第二卷　第一分册

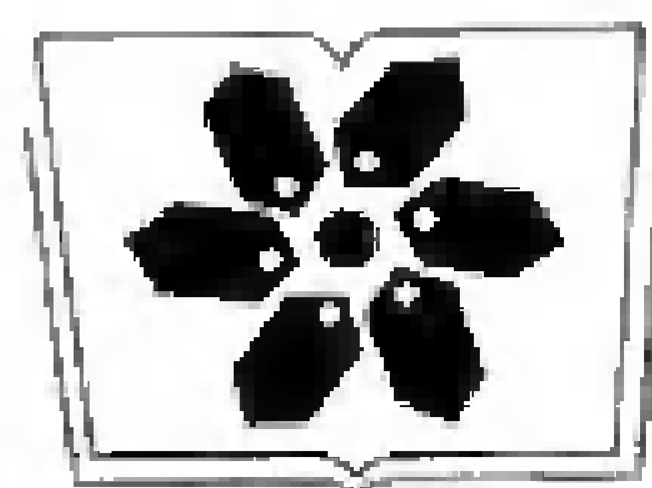
数学名著译丛

微积分和
数学分析引论

R.柯朗　F.约翰　著



科学出版社
www.sciencep.com



中国科学院科学出版基金资助出版

数学名著译丛

微积分和数学分析引论

第二卷 第一分册

R. 柯朗 F. 约翰 著

林建祥 刘婉如 朱德威 等 译

冷生明 校

科学出版社

2001

图字：01-98-2682

内 容 简 介

本书系统地阐述了微积分学的基本理论。在叙述上，作者尽量作到既严谨而又通俗易懂，并指出概念之间的内在联系和直观背景。原书分两卷，第一卷为单变量情形，第二卷为多变量情形。

第二卷中译本分为两册出版。本书是第二卷第一分册，包括前三章。第一章详论多元函数及其导数，包括线性微分型及其积分，补充了数学分析中最基本的概念的严密证明；第二章在线性代数方面为现代数学分析的基础准备了充分的材料；第三章叙述多元微分学的发展及应用，包括隐函数存在定理的严密证明，多元变换与映射的基本理论，曲线、曲面的微分几何基础知识以及外微分型等基本概念。原书有练习解答，分别编入各分册。

译者（按内容顺序）：邵士敏、周建莹、张锦炎（第一章）、刘婉如（第二章）、林建祥、张顺燕、朱德威（第三章）、林源渠（解答）。

读者对象为高等学校理工科师生与工程技术人员。

图书在版编目（CIP）数据

微积分和数学分析引论．第二卷，第一分册 /【美】R. 柯朗，【美】F. 约翰著；林建祥等译．——北京：科学出版社，2001.3
书名原名：Introduction to Calculus and Analysis, Volume 2
ISBN 7-03-008540-X

I. 微… II. ①柯…②约…③林… III. ①微积分②数学分析 IV. O17

中国版本图书馆数据核字（2000）第60214号

科学出版社 出版
北京东黄城根北街16号
邮政编码：100717
北京双青印刷厂 印刷
科学出版社发行 各地新华书店经销
*
2001年3月第 一 版 开本：850 × 1168 1/32
2001年3月第一次印刷 印张：16
印数：1-3 000 字数：416 000

定价：32.00元

（如有印装质量问题，我社负责调换（环伟））

序　　言

R. 柯朗的《微积分 (Differential and Integral Calculus)》，卷一和卷二，获得了巨大的成功，引导了几代数学家进入高等数学的领域。整套书阐发了这样一条富有教益的道理：真正有意义的数学是由直观想象与演绎推理相结合而创造出来的. 在准备本修订版的过程中，著者们力图保持原著所特有的这两种思维方式的合理结合. 虽然 R. 柯朗未能亲眼看到这第二卷的修订本的出版，但是在柯朗博士于 1972 年 1 月去世之前，所有主要的修改内容都已经由著者们商量出一致的意见并草拟了纲要.

从一开始，著者们就清楚地理解到，阐述多元函数的第二卷应当比第一卷作更大的修改. 特别是，似乎最好用与一维空间中的积分法同样程度的严密性和普遍性来处理高维空间中的积分法的所有基本定理. 另外，若干具有根本重要性的新概念和新论题，它们在著者们看来，是属于数学分析引论的.

本卷末尾较短的几章 (六、七、八) (分别讲微分方程、变分法、复变函数的) 仅作了较小的改动. 在本卷的主体部分 (第一至第五章) 中，我们尽量保存了原来的体系，各章的主题都在不同的水平上按大体上平行的两条线索进行阐述：一条是较多地基于直观论证的非正式引述，一条是为后续的严密证明打基础的、对于应用的讨论.

原来的第一章中所讲的线性代数的材料，看来已不足以作为扩展了的微积分结构的基础. 因此，这一章 (现在的第二章) 完全重写了，现在讲述了所需要的 n 阶的行列式、矩阵、多线性型、格拉姆行列式、线性流形等的一切性质.

现在的第一章包括了线性微分型及其积分的所有基本性质. 这就为读者提供了阅读第三章 (第 3.6 节) 新添的高阶外微分型的预备

知识. 现在的第三章中还新添了一个关于隐函数定理的利用逐次逼近法的证明, 一个关于临界点的个数和二维向量场的指数的讨论.

在第四和第五章中, 对重积分的基本性质作了大量的增补. 这里面临着一个大家熟知的困难: 展布在一个流形 M 上的积分, 很容易通过把 M 适当划分成小片来定义, 却必需证明它不依赖于特殊的分划. 这由系统地使用约当可测集类的有限交性质和单位分划而解决了. 为了使应用拓扑学上的复杂性减少到最低限度, 我们只考虑了光滑地嵌入于欧氏空间中的流形. 另外, 对流形的 "定向" 的概念也作了详细的研讨, 这是要讨论外微分型的积分及其可加性所必需的. 在这个基础上我们给出了 n 维空间中的散度定理和司铎克斯定理的证明. 在第四章关于傅里叶积分的一节 (第 4.13 节) 里, 还新添了关于帕塞瓦尔等式和傅里叶重积分的论述.

在这一卷的准备过程中, 最珍贵的是, 著者们能不断得到两位朋友慷慨的帮助, 一位是 Carnegie-Mellon 大学的 Albert A. Blank 教授, 一位是 Negev 大学的 Alan Solomon 教授. 在几乎每一页上都有着他们的批评、改正、建议的痕迹. 他们还为这一卷准备了练习题和问题.

我们感谢 K.O. Friedrichs 教授和 Donald Ludwig 教授建设性的宝贵建议, 还感谢 John Wiley and Sons 公司及其编辑部的不断鼓励和帮助.

F. 约翰

纽约, 1973 年 9 月

目　　录

第一章　多元函数及其导数

在第一卷中曾讨论过的极限、连续、导数和积分等概念，同样也是二元或多元函数的基本概念. 然而，有很多在一元函数理论中并不存在的新的现象，必须在多维中加以讨论. 通常一个定理只要对于两个变量的函数可以证明它，那么在证明中不需要作任何本质的改变，就容易推广到多于两个变量的函数中去. 因此，在以后的论述中我们常限于讨论两个变量的函数，其中各种关系都比较容易用几何图形来显示，而只当由此得到一些另外的见解时，才对三个或更多个变量的函数加以讨论；所得结果也同样可以作简单的几何学的解释.

1.1　平面和空间的点和点集

a. 点的序列：收敛性

在一个笛卡儿平面坐标系中，一对有顺序的数值 (x, y) 在几何上可以用一个点 P 来表示，这个点的坐标为 x 和 y. 两点 $P=(x, y)$ 与 $P'=(x', y')$ 之间的距离可以由公式

$$\overline{PP'}=\sqrt{(x'-x)^2+(y'-y)^2}$$

求得. 这是欧几里得几何学的基本公式. 我们利用距离的概念可以定义一个点的邻域. 一个点 $C=(\alpha, \beta)$ 的 ε 邻域是由所有那些与 C 的距离小于 ε 的点 $P=(x, y)$ 构成的；从几何学上说，这是一个

以 C 为中心、以 ε 为半径的圆盘[1)]，它可用不等式

$$(x-\alpha)^2+(y-\beta)^2<\varepsilon$$

来描述.

我们考虑无穷的点序列

$$P_1=(x_1,y_1),P_2=(x_2,y_2),\cdots,P_n=(x_n,y_n),\cdots.$$

例如 $P_n=(n,n^2)$ 定义了一个序列，其中所有的点都在抛物线 $y=x^2$ 上. 一序列中的点并不一定都不相同. 例如，无穷序列 $P_n=(2,(-1)^n)$ 只有两个不同的元素.

如果能找到一个圆盘，它包含所有的 P_n，也就是说，如果存在一个点 Q 与一个数 M，使得对所有的 n 都有 $\overline{P_nQ}<M$，那么我们说序列 $P_1,P_2,\cdots$ 是**有界的**. 例如序列 $P_n=(1/n,1/n^2)$ 是有界的，而序列 (n,n^2) 是无界的.

与序列有关的最重要的概念是**收敛**的概念. 我们说，一个点序列 $P_1,P_2,\cdots$ 收敛到一个点 Q，或

$$\lim_{n\to\infty}P_n=Q,$$

是指距离 $\overline{P_nQ}$ 收敛到零. 这样，$\lim\limits_{n\to\infty}P_n=Q$ 就意味着对于每一个 $\varepsilon>0$，都必然存在一个数 N，使对于所有 $n>N$，P_n 都在 Q 的 ε 邻域内[2)].

举一个例. 对于由 $P_n=(e^{-n/4}\cos n,e^{-n/4}\sin n)$ 定义的点序列，我们有 $\lim\limits_{n\to\infty}P_n=(0,0)=Q$，因为，在这里

$$\overline{P_nQ}=e^{-n/4}\to 0\quad 当\ n\to\infty.$$

---

1) 平常所用的"圆"这个词，是指一条曲线还是指由它所界的区域，是含糊不清的. 我们根据实际流行的说法，把"圆"只用于曲线，而把"圆形区域"或"圆盘"用于两维区域. 同样，在空间中我们把"球面"(即球形曲面) 与它所界的三维立体"球体"区别开来.

2) 等价地说，任何一个以 Q 为中心的圆盘，除有限个 P_n 外，它包含所有的 P_n. 我们也常用记法表示成：当 $n\to\infty$ 时 $P_n\to Q$.

我们指出，P_n 是沿着极坐标方程为 $r = e^{-\theta/4}$ 的对数螺线趋于原点 Q 的 (见图 1.1).

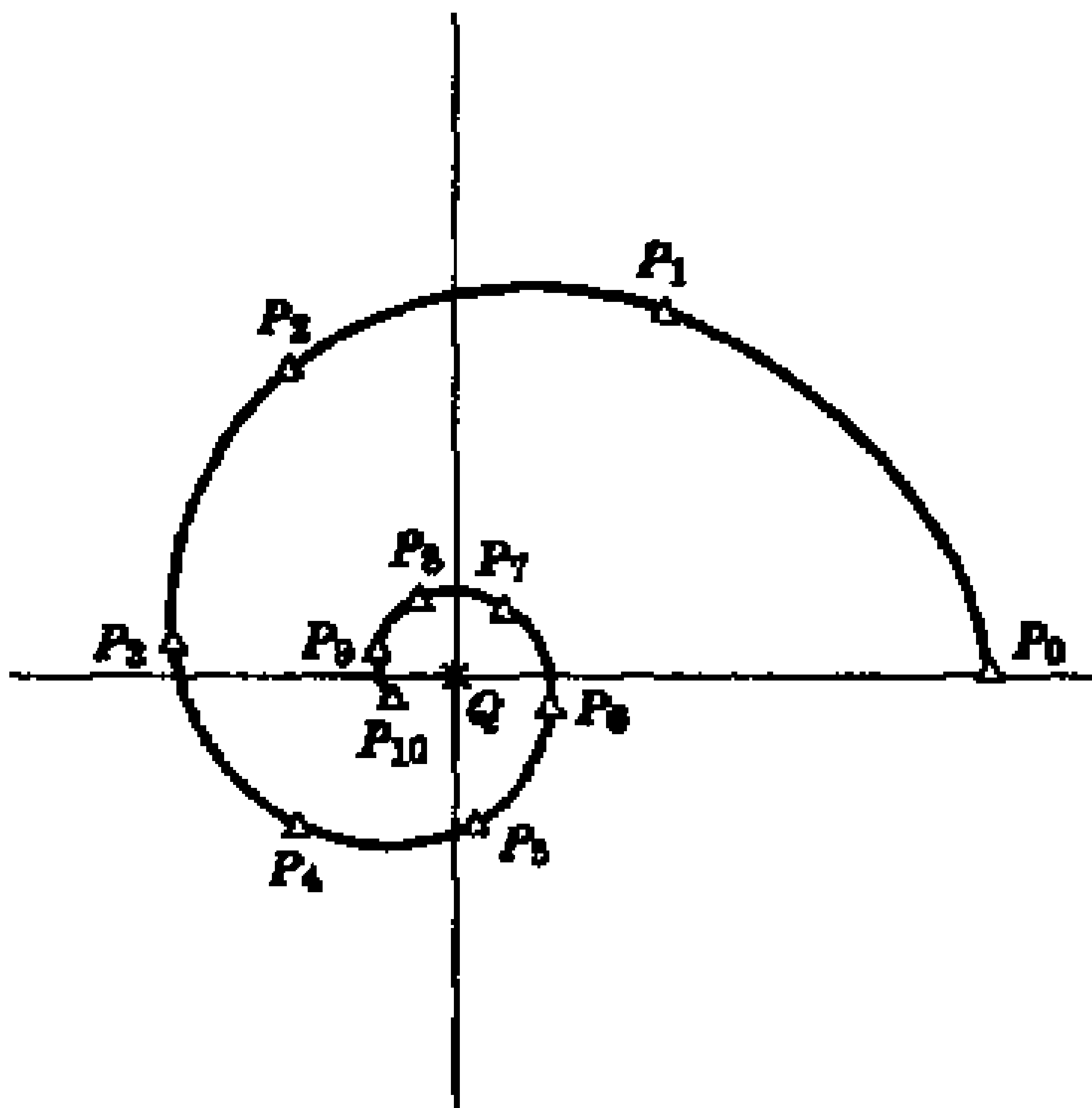


图 1.1　收敛序列 P_n

点序列 $P_n = (x_n, y_n)$ 收敛到点 $Q = (a, b)$ 意味着，两个数序列 x_n 与 y_n 分别收敛，且

$$\lim_{n\to\infty} x_n = a, \quad \lim_{n\to\infty} y_n = b.$$

诚然，$\overline{P_nQ}$ 很小隐含着 $x_n - a$ 与 $y_n - b$ 都很小，因为

$$|x_n - a| \le \overline{P_nQ}, \quad |y_n - b| \le \overline{P_nQ};$$

反之

$$\overline{P_nQ} = \sqrt{(x_n - a)^2 + (y_n - b)^2} \le |x_n - a| + |y_n - b|,$$

从而当 $x_n \to a$ 与 $y_n \to b$ 时有 $\overline{P_nQ} \to 0$.

如同数序列的情形一样，我们可以利用 柯西的内在收敛判别法 来证明一个点序列收敛，而不需要知道它的极限值. 这个判别

法，在两维中断言：一个点序列 $P_n=(x_n,y_n)$ 是收敛的必要充分条件是，对每一个 $\epsilon>0$，不等式 $\overline{P_nP_m}<\epsilon$ 对大于适当的值 $N=N(\epsilon)$ 的所有 n,m 都成立．其证明可以对两个序列 x_n 与 y_n 中的每一个运用数序列的柯西收敛判别法来推得．

b. 平面上的点集

在讨论单变量 x 的函数时，我们常允许 x 在一个"区间"内变化，区间可以是闭的或开的，可以是有界的或无界的．而对于高维空间中的函数所可能取的区域而言，必需考虑更多种类的集合，并且必需引进描述这些种类集合的简单性质的一些述语．在平面中我们通常考虑的可以是曲线也可以是二维区域．平面曲线在第一卷第四章中已广泛地讨论过．通常它们可以用"非参数"形式的一个函数 $y=f(x)$ 给出，或者用"参数"形式的一对函数 $x=\phi(t)$, $y=\psi(t)$ 给出，或者用一个隐式方程 $F(x,y)=0$ 给出（在第三章中我们将更多地讲到隐式表示法）．

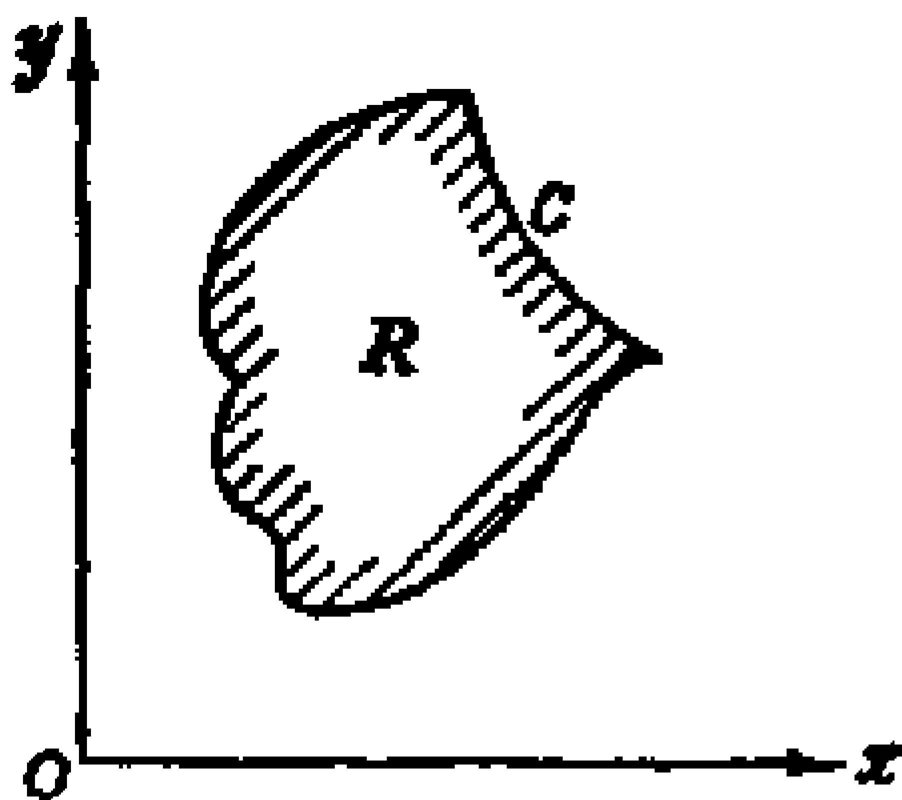


图 1.2　单连通区域

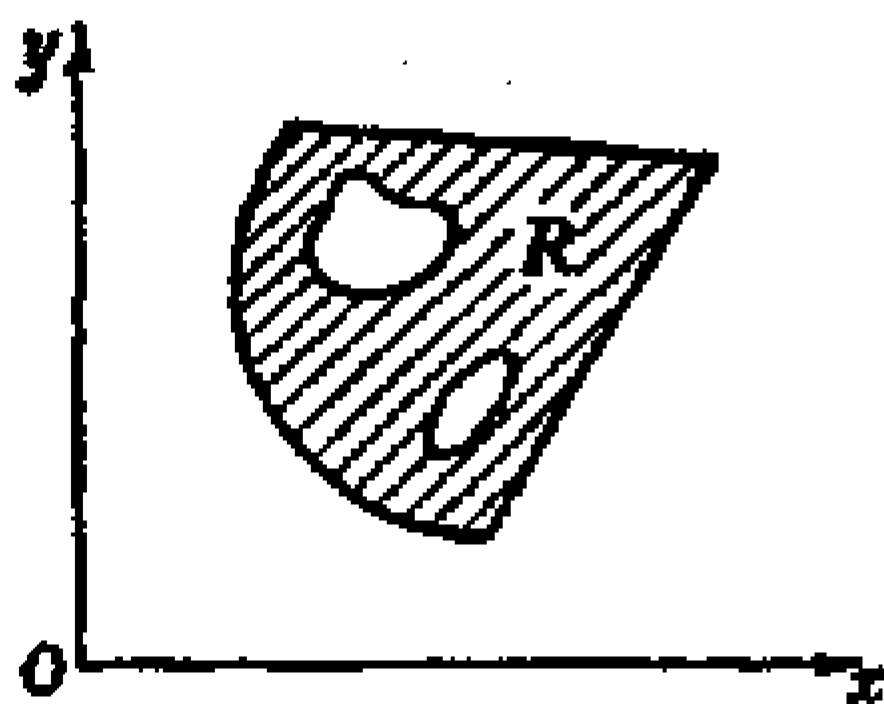


图 1.3　三连通区域

除曲线之外，我们还有组成一个**区域**的**二维**点集．这个区域可以是整个 xy 平面，或者是由一简单闭曲线围成的一部分平面（在这种情况下，形成一个**单连通**区域如图 1.2 所示），或者是由几个这类曲线围成的一部分平面．在后一种情况下，我们称之为**多连通**区域，边界曲线的数目就叫做**连通数**；例如图 1.3 就表示一

个 三连通 区域. 一个平面集合也可以是全然不连通的[1], 而是由几个分离部分组成的 (图 1.4).

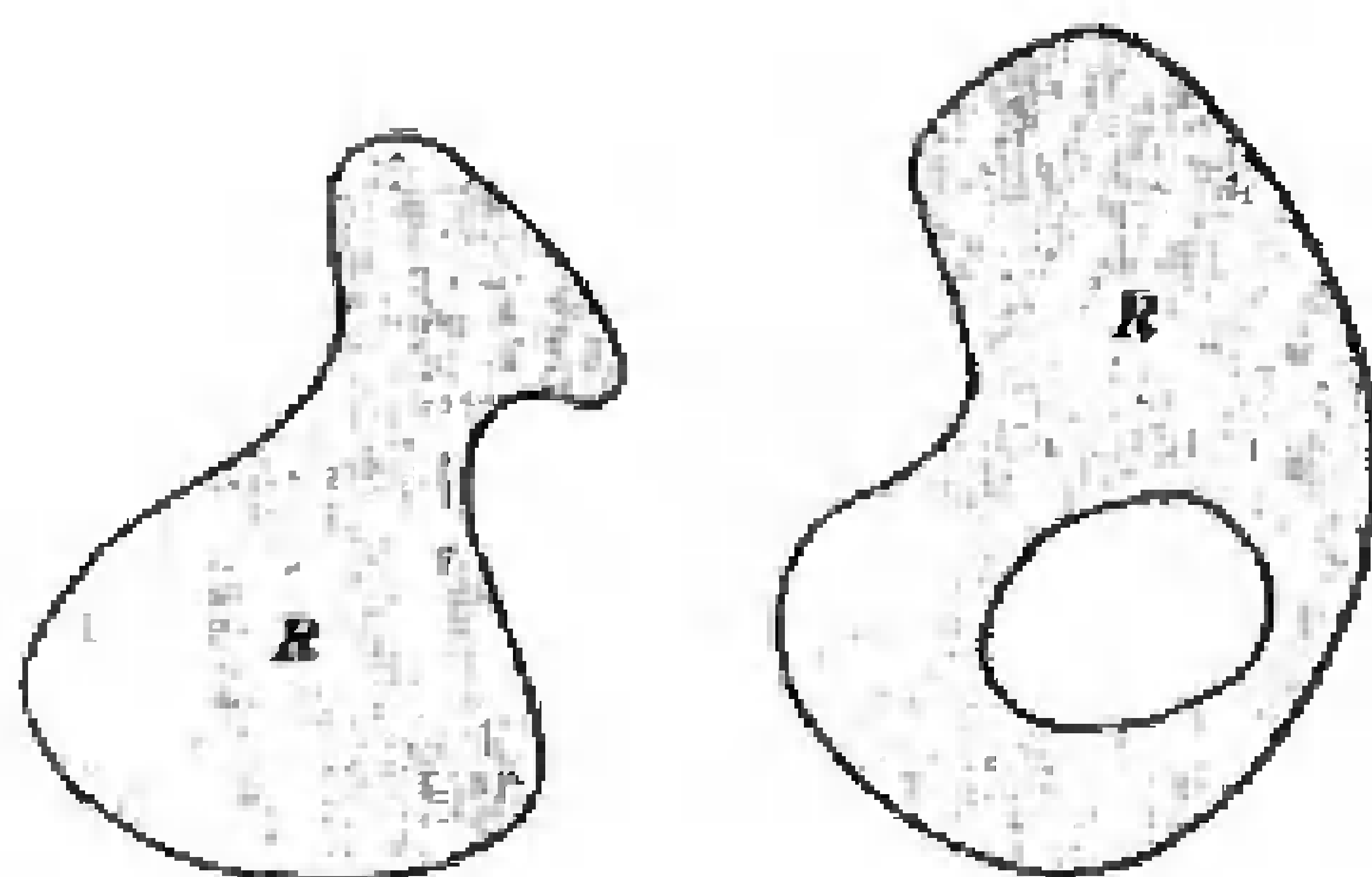


图 1.4　非连通区域

一般说来, 要考虑的区域的边界曲线都是 逐段光滑 的, 亦即每一个这样的曲线是由有限个弧组成的, 每一段弧上所有的点, 包括端点在内, 都有一个连续转动的切线. 因此, 这样的曲线至多也只有有限个角.

在绝大多数情况下, 我们用一个或多个不等式来描述一个区域, 而在边界的一些部分保持等号. 有两种最重要的区域形式是经常遇到的: 一个是矩形区域 (其各边平行于坐标轴), 一个是圆盘. 矩形区域(图 1.5) 是由这样一些点 (x,y) 构成的, 它们的坐标满足不等式

$$a < x < b, \quad c < y < d;$$

每一个坐标限制在一个确定的区间上, 并且点 (x,y) 在一个矩形内部变化. 我们这里定义的矩形区域是 开的, 就是说它不包含它的边界.

把定义该区域的一个或几个不等式改为等式, 并且允许 (但不是必需的) 其余的不等式中有等号, 这样就得到边界曲线. 例如

$$x = a, \quad c \leq y \leq d$$

1) "连通" 的确切定义见第 110 页.

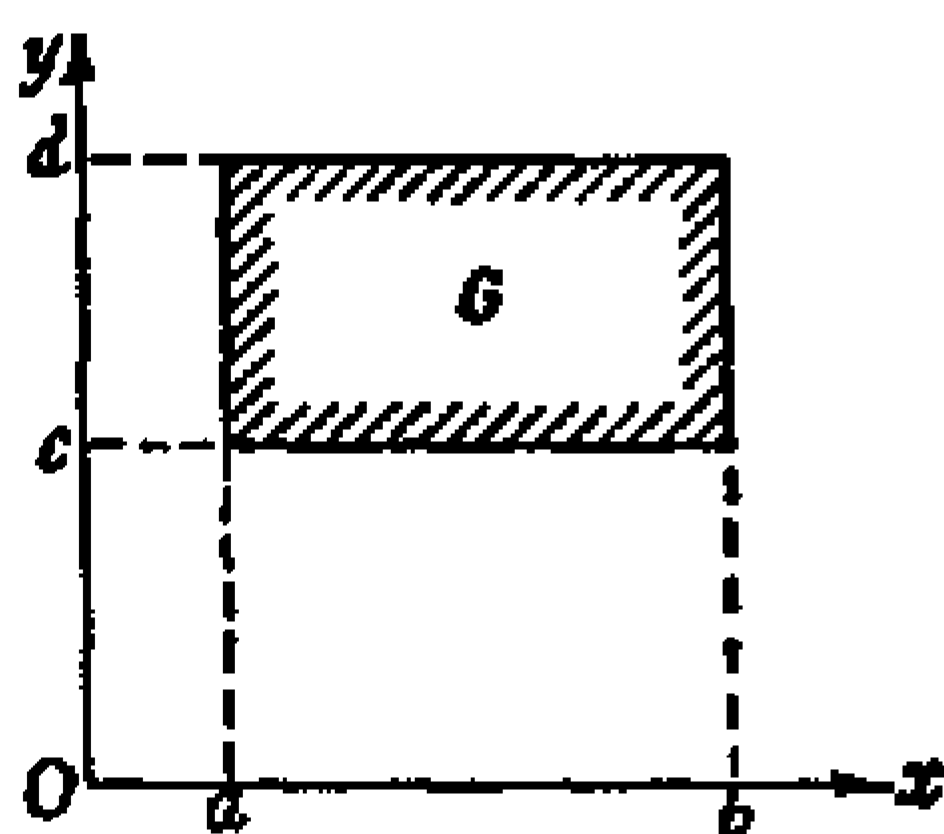


图 1.5 矩形区域

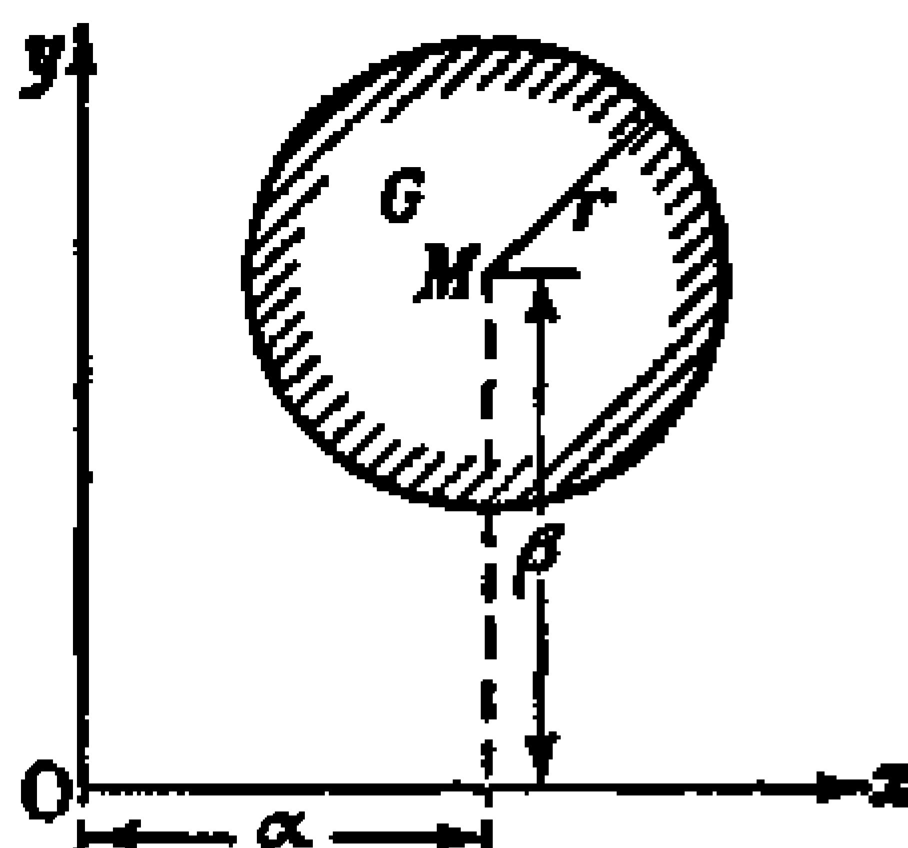


图 1.6 圆盘

定义了矩形的一条边. 把所有的边界点加到该集合中去, 就得到 **闭矩形**, 它由下面的不等式来描述:

$$a \leq x \leq b, \quad c \leq y \leq d.$$

如前面所见那样, 中心为 (α,β)、半径为 r 的 **圆盘**(图 1.6), 可由下面这不等式给出:

$$(x-\alpha)^2+(y-\beta)^2<r^2.$$

把边界圆加到这个 "开" 圆盘上, 我们就得到 "闭" 圆盘, 它由下式表示

$$(x-\alpha)^2+(y-\beta)^2\leq r^2.$$

c. 集合的边界. 闭集与开集

区域的边界可以看作一类薄膜, 它把那些属于区域的点与不属于区域的点分开. 我们将会看到, 这样的边界的直观概念并不总是有意义的. 然而, 值得注意的是, 有一种方法可以十分一般地定义任何点集的 **边界**, 使得它在这种定义方式中至少与我们的直观概念相一致. 我们说 **点 P 是点集合 S 的一个边界点, 意思是说, P 的每一个邻域内有属于 S 的点也有不属于 S 的点**. 因此, 如果 P 并不是一个边界点, 那就必定存在 P 的一个邻域, 它只包含同一

类的点；也就是说，或者我们可以找到 P 的一个邻域，它包含的全部是 S 的点，在这种情况下，我们叫 P 为 S 的一个**内点**；或者我们可以找到 P 的一个邻域，它包含的全部不是 S 的点，在这种情况下，我们叫 P 为 S 的一个**外点**. 这样，**对于一个给定的点集 S，平面上的每一个点，不是 S 的边界点，就是 S 的内点或外点，而且只能属于这三类点中的一类**. S 的全部边界点的集合构成 S 的**边界**，我们用符号 ∂S 表示.

举一个例子：令 S 为矩形区域

$$a < x < b, \quad c < y < d.$$

显然，对于 S 中任一个点 P，我们可以找到一个以 $P = (\alpha, \beta)$ 为中心的小圆盘，它全部包含于 S 之中；我们只需要取一个 P 的 ε 邻域，其中 ε 为足够小的正值，使

$$a < \alpha - \varepsilon < \alpha + \varepsilon < b, \quad c < \beta - \varepsilon < \beta + \varepsilon < d.$$

这表明在这里 S 的每一个点都是内点. S 的边界点 P 是那些刚好位于矩形的一个边上或一个角上的点. 在第一种情况下，P 的每一个足够小的邻域一半属于 S，一半不属于 S；在第二种情况下，每一个邻域的四分之一属于 S 而四分之三不属于 S (图 1.7).

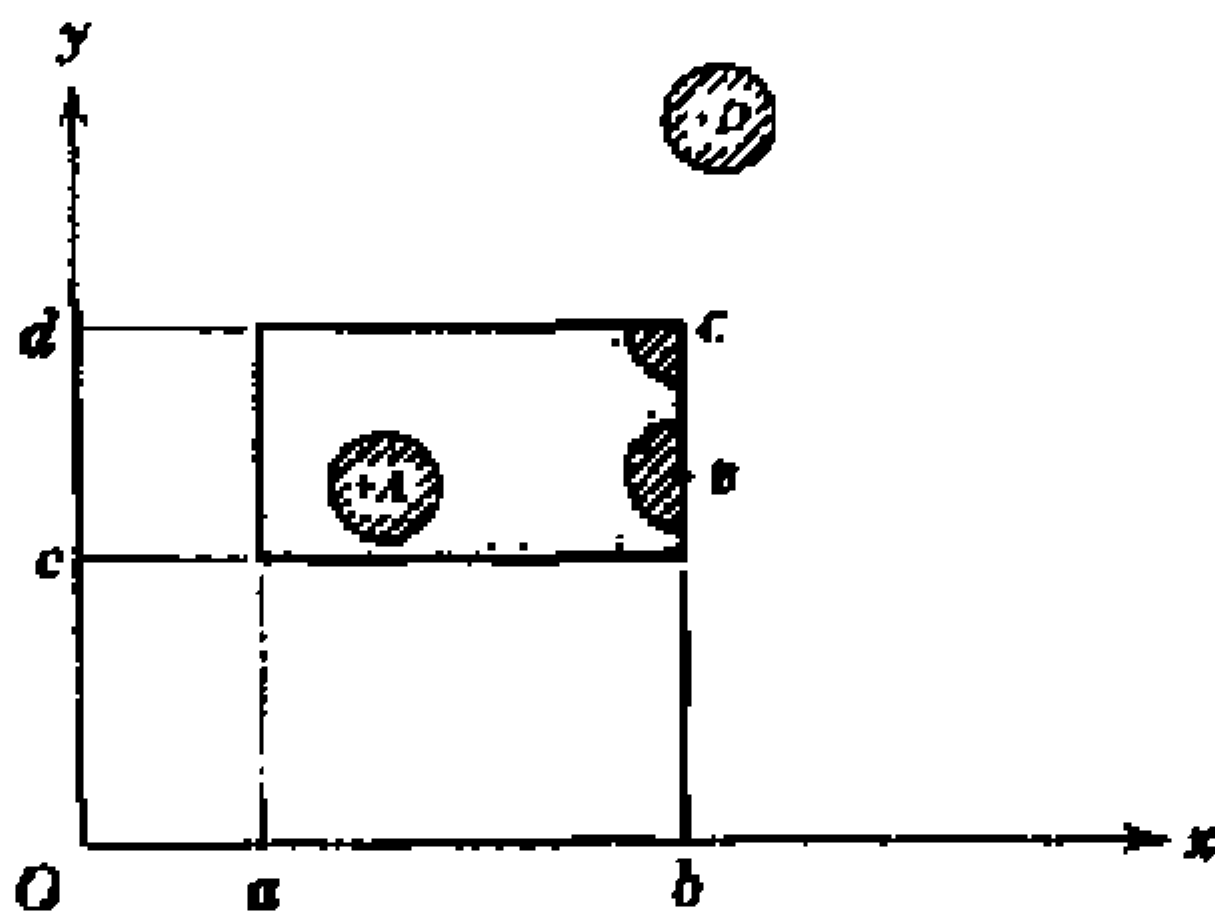


图 1.7 矩形区域的内点 A，外点 D 和边界点 B, C

根据定义，集合 S 的每一个内点 P 必须是 S 的点，这是因为存在 P 的一个邻域，它的点全部都是 S 的点，而 P 又属于这个邻域．同样，S 的任何一个外点明确地不属于 S．另一方面，一个集合的边界上的点，有时属于有时不属于这集合[1]．开矩形

$$a < x < b, \quad c < y < d$$

并不包含其边界点，而闭矩形

$$a \le x \le b, \quad c \le y \le d$$

则包含其边界点．

一般说来，我们称一个点集 S 是**开**的，如果 S 的边界点都不属于 S（亦即，如果 S 由其全部内点组成）．S 叫做**闭**的，如果它包含它的边界．对于任何一个集合 S，我们总可以把一切原先并不属于 S 的那些边界点加到 S 上而得到一个闭集．这样我们就得到一个新的集合，称为 S 的**闭包** $\overline{S}$．读者可以容易地证明，S 的闭包是一个闭集．外点是那些不属于 S 的闭包的点．同样地，我们可以定义 S 的内点组成的集合为 S 的**内部** S^0，这个集合可以从 S 中减去全部边界点而得到．S 的内部是开的．

必须看到，集合可以既不是开的也不是闭的．我们可以容易地构造一个集合 S，它只包含它的边界的一部分，例如半开矩形

$$a \le x < b, \quad c \le y < d.$$

把我们对边界的观念应用到很一般的集合中去，并将结果从直觉中解脱出来，这也是一件非常重要的事．最好的例子是由平面上全部“有理点”构成的集合 S，即由那些坐标 x, y 都是有理数的点 $P=(x, y)$ 构成的集合，把它说成是一条“曲线”或一个“区域”都无意义．很清楚，每一个平面上的圆盘都包含有理点和无理点．因

1) 注意“不属于 S”和“S 之外”是有区别的．S 的一个边界点即使不属于 S 也绝不可能在 S 之外．

此，这里没有边界曲线；它的边界 ∂S 包含整个平面，它既无内点也无外点.

即使在边界是一维的情况，也并不是所有边界点都是把内点与外点分隔开. 例如，不等式

$$(x-\alpha)^2+(y-\beta)^2<r^2,\quad y\neq\beta$$

描述了把一个直径除外的圆盘；这里，边界系由圆

$$(x-\alpha)^2+(y-\beta)^2=r^2$$

与直径

$$y=\beta,\quad |x-\alpha|<r$$

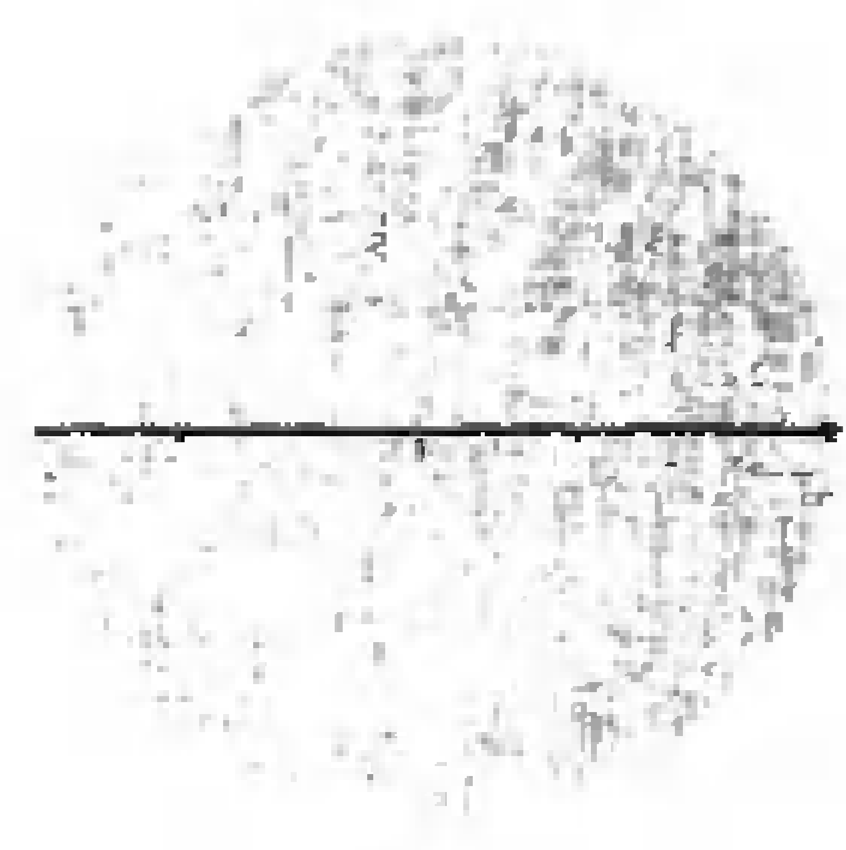


图 1.8　除去了一个直径的圆盘

组成. 而直径上每一个 (非端点的) 点的任何足够小的邻域都不包含外点 (图 1.8).

d. 闭包作为极限点的集合

当我们考虑那些全部属于集合 S 的点序列 $P_1, P_2, \cdots$ 的极限时，一个集合的 "内部"、"边界" 和 "外部" 的概念是十分重要的 [1]. 很显然，S 的外部的一点 Q 不能成为序列的极限点，这是因为存在 Q 的一个邻域，它没有 S 的点，这就阻止了 P_K 任意接近 Q. 因此 S 内的点序列的极限必须是 S 的边界点或内点. 由于 S 的内部与全部边界点组成 S 的闭包，由此推知，S 内的序列的极限属于 S 的闭包.

反之，S 的闭包的每一个点 Q 实际上都是 S 的某一序列 $P_1, P_2, \cdots$ 的极限. 因为如果 Q 是闭包的一个点，则 Q 或者属于 S 或者属于边界. 在第一种情形，我们显然有 S 的一个点序列 $Q, Q, Q, \cdots$

---

1) 这些点 P_K 不必彼此之间不相同.

收敛到 Q. 在第二种情形，对于任何 $\varepsilon > 0$, Q 的 ε 邻域都至少包含一个 S 的点. 对于每一个自然数 n, 我们可以在 S 中选择一个点 P_n, 它属于 Q 的 ε 邻域，而

$$\varepsilon = \frac{1}{n}.$$

很清楚，P_n 收敛到 Q.

e. 空间的点与点集

一组有序的三个数 (x, y, z) 可用通常的方式由空间中的一个点 P 来表示. 这里数 x, y, z 为 P 的笛卡儿坐标，是从 P 到三个互相垂直的平面的 (带有符号的) 距离. 两个点 $P(x, y, z)$ 与 $P'(x', y', z')$ 之间的距离 $\overline{PP'}$ 由下式给出:

$$\overline{PP'} = \sqrt{(x'-x)^2 + (y'-y)^2 + (z'-z)^2}.$$

点 $Q = (a, b, c)$ 的 ε 邻域是由那些适合 $\overline{PQ} < \varepsilon$ 的点 $P = (x, y, z)$ 构成的；这些点构成由不等式

$$(x-a)^2 + (y-b)^2 + (z-c)^2 < \varepsilon^2.$$

给出的球.

与平面上矩形区域相类似的是平行六面体 [1], 由下列一组不等式所描述:

$$a < x < b, \quad c < y < d, \quad e < z < f.$$

所有为平面集合而发展的概念 —— 边界、闭包等等 —— 都可用一种明显的方式推广到三维中去.

当我们处理有序的四个数 x, y, z, w 所成的数组时，我们的直观就无法提供一个几何解释. 但是我们仍然可以方便地使用几何术语，称呼 (x, y, z, w) 为一个 "四维空间中的点". 满足不等式

$$(x-a)^2 + (y-b)^2 + (z-c)^2 + (w-d)^2 < \varepsilon^2$$

1) 平行的 epipedon (希腊语 "平面").

的那些四维数组 (x,y,z,w) 就按定义构成点 (a,b,c,d) 的 ϵ 邻域.

一个矩形区域[1] 用如下形式的一组不等式描述:

$$a<x<b,\quad c<y<d,\quad e<z<f,\quad g<w<h.$$

当然，这种四维中的"点"的概念并不神秘，它仅仅是一个方便的术语，并不意味着四维空间的物理实体．诚然也并没有什么东西能阻止我们把 n 维数组 $(x_1,x_2,\cdots,x_n)$ 叫做 n 维空间中的一个"点"，这里 n 可以是任何一个自然数．在许多应用中，通过这种方式把 n 个量所描述的一个系统用高维空间的一个单个点来代表，这办法是既很有用也很有启发性的[2]．人们常常模拟三维空间的几何解释，作为探讨高于三维时的指导.

练　习　1.1

1. 平面上一个点 (x,y) 可以用一个形如 $Z=x+iy$ 的一个复数表示 (见第一卷第 114 页). 试讨论下列取不同值的 Z 序列的收敛性:

(a) Z^n;

(b) $Z^{1/n}$, 这里 $Z^{1/n}$ 定义为 Z 的 n 次元根，即最小正幅角的根.

2. 证明，对 $P_n=(x_n+\xi_n,y_n+\eta_n)$, 则有

$$\lim_{n\to\infty}P_n=(x+\xi,y+\eta),$$

这里，假设了极限 $x=\lim\limits_{n\to\infty}x_n$, $\xi=\lim\limits_{n\to\infty}\xi_n$, $y=\lim\limits_{n\to\infty}y_n$, $\eta=\lim\limits_{n\to\infty}\eta_n$ 都存在.

3. 证明圆盘 $x^2+y^2<1$ 的每一个点均为内点. 这对于 $x^2+y^2\le 1$ 是否也对? 作出解释.

---

1) 在多维中这类矩形区域也常用"单元"和"区间"这种名词来描述.

2) 例如一个容器中气体的分子所成的系统可以用一个具有很高维数的"相位空间"中的一个单个的点的位置来描述. 再说得广泛一些，在分析的某些部分中习惯于将数 $x_1,x_2,\cdots$ 所构成的一个无穷序列用一个无穷维空间中的一个点来代表.

4. 证明适合条件 $y > x^2$ 的全部点 (x, y) 所成的集合 S 是开的.

5. 当把一个线段考虑成 xy 平面的一个子集时，什么是这个线段的边界？

问　题　1.1

1. 设 P 为集合 S 的一个边界点，但不属于 S. 试证明 S 中存在一 互不相同的点 的序列 $P_1, P_1, \cdots$, 以 P 为极限.

2. 试证一个集的闭包是闭的.

3. 设 P 为集合 S 中的任何一个点，并设 Q 为这集合的任何一个外点. 试证线段 PQ 包含 S 的一个边界点.

4. 设 G 为满足条件 $|x| < 1$, $|y| < \frac{1}{2}$, 以及 $x = \frac{1}{2}$, $y < 0$ 的全部点 (x, y) 所成的集合. 问 G 是否只包含内点？试证明之.

1.2　几个自变量的函数

a. 函数及其定义域

这样形式的方程

$$u = x + y, \quad u = x^2y^2 \text{ 或 } u = \log(1 - x^2 - y^2)$$

对于每一对值 (x, y) 都确定一个 函数值 u. 在前两个例子中，对每一对值 (x, y) 都确定了 u 的一个值，而在第三个例子中，只有当一对值 (x, y) 满足不等式 $x^2 + y^2 < 1$ 时方程才有意义.

一般说来，对于每一对属于某一个特定集合 (即函数的 定义域) 的 (x, y) 值，依照某规律 f, 确定了 因变量 u 的唯一的一个值，我们就说 u 是 自变量 x 和 y 的一个函数. 这样，一个函数 $u = f(x, y)$ 便定义了一个 映射，把 xy 平面上的一个点集 (即 f 的

定义域) 映射到 u 轴上的某一点集，即 f 的值域. 类似地，我们说 u 是 n 个变量 $x_1, x_2, \cdots, x_n$ 的一个函数，是指对于某一特定集合中的每一组值 $(x_1, x_2, \cdots, x_n)$ 都确定了一个相应的唯一值 u.[1)]

例如，一个直平行六面体的体积 $u = xyz$ 是三个边长 x, y, z 的一个函数；磁倾角是纬度、经度和时间的一个函数；和 $x_1 + x_2 + \cdots + x_n$ 是 n 个项 $x_1, x_2, \cdots, x_n$ 的一个函数.

必须指出：一个函数 f 的定义域是描述它的一个不可缺少的部分. 在 $u = f(x, y)$ 由显示式给出的情形，自然要取能使这表示式有意义的所有 (x, y) 作为 f 的定义域. 然而，用同一表示式所给出的函数，可以用"限制"的方法来定义一个较小的域，公式 $u = x^2 + y^2$ 可以用来定义一个具有定义域

$$x^2 + y^2 \leq \frac{1}{2}$$

的函数.

像一元函数的情形一样，一个函数关系 $u = f(x, y)$ 把自变量组 x, y 联系到 u 的一个唯一值. 因此，一个函数值不能由多值的解析表示式如像 $\arctan \frac{y}{x}$ 所确定，除非我们规定 "arc tangent" 只取主分枝，即其值在 $-\frac{\pi}{2}$ 到 $+\frac{\pi}{2}$ 之间 (参阅第一卷第 240 页)，并且我们还要除去直线 $x = 0$.[2)]

b. 最简单的函数

如同单变量的情形一样，多于一个变量的最简单的函数仍是有理整函数或多项式. 最一般的一次多项式或线性函数具有这样的

---

1) 我们常常宁愿把函数 f 看作是给一个点 P 指定一个值，而不是给描写 P 的一对坐标 (x, y) 指定一个值. 这样我们就把 $f(x, y)$ 写作 $f(P)$. 这种记号，当点 P 与值 $f(P)$ 之间的函数关系由几何来定义而不是参考一个特定的 xy 坐标系时，特别适用.

2) 取主值，我们可看到 $u = \arctan y/x$，当 $x>0$，不是别的，而是从正 x 轴算起，点 (x, y) 的极角. 这个极角还可以用明显的方式几何地定义为一个单值函数 (取值于 $-\pi$ 到 π 之间)，如果我们除去原点以及位于负 x 轴上的点；但如果我们把 arctangent 理解为主分枝，那么在推广了的区域内这个极角就不再由 $\arctan y/x$ 给出来了.

形式：

$$u = ax + by + c,$$

其中 a, b 和 c 均为常数. 最一般的二次多项式具有这样的形式：

$$u = ax^2 + bxy + cy^2 + dx + ey + f,$$

它的定义域是整个 xy 平面. 任何次数的一般多项式是有限个项 $a_{mn}x^m y^n$ (叫做单项式) 的和，其中 m 和 n 为非负整数，而系数 a_{mn} 是任意的.

只要系数 a_{mn} 不为 0, 单项式 $a_{mn}x^m y^n$ 的次数是 x 和 y 的表示式中的和 $m+n$. 多项式的次数是 (在相同幂的 x 和 y 合并项以后) 系数不为 0 的任何一个单项式的最高次数. 一个多项式，其所有各单项式具有相同的次数 N 时叫做一个齐次多项式或 N 次型. 如 $x^2 + 2xy$ 或 $3x^3 + (7/5)x^2 y + 2y^3$ 就是这种型.

通过有理函数开方，我们可得到一些代数函数[1]，例如

$$u = \sqrt{\frac{x-y}{x+y}} + \sqrt[3]{\frac{(x+y)^2}{x^3+xy}}.$$

我们将要用到的大多数更复杂的多变量函数，都可以用单变量的那些熟悉的函数来描述，例如

$$u = \sin(x \arccos y) \text{ 或 } u = \log_x y.$$

c. 函数的几何表示法

如同我们用曲线来表示一元函数一样，我们可用曲面来几何地表示二元函数. 为了这个目的，我们来考虑一个空间的 (x, y, u) 直角坐标系，并且在 xy 平面上的函数定义域 R 内的每一个点 (x, y) 上方，标出以 $u = f(x, y)$ 为第三坐标的点 P. 当点 (x, y) 遍及定义

1)“代数函数”一词的一般定义见第 249 页.

域 R 时，点 P 描绘出一个空间曲面，这个曲面我们就取作为函数的几何表示.

反过来说，在解析几何中，空间曲面可用二元函数表示，因此在这种曲面与二元函数之间有一种互逆关系．例如，函数

$$u=\sqrt{1-x^2-y^2}$$

对应着位于 xy 平面之上的半球，球半径为 1，球心在原点．函数 $u=x^2+y^2$ 对应着一个所谓**旋转抛物面**，它是把一个抛物线 $u=x^2$ 绕 u 轴旋转而得到的 (图 1.9)．函数 $u=x^2-y^2$ 和 $u=xy$ 对应着双曲抛物面 (图 1.10)．线性函数 $u=ax+by+c$ 的图形是一个空间平面．如果在函数 $u=f(x,y)$ 中一个自变量，譬如 y，并不出现，因而 u 只依赖于 x，即 $u=g(x)$，那么这函数就在 x,y,u 空间中代表一个柱面，它是由通过曲线 $u=g(x)$ 上的点且与 ux 平面垂直的直线所产生的.

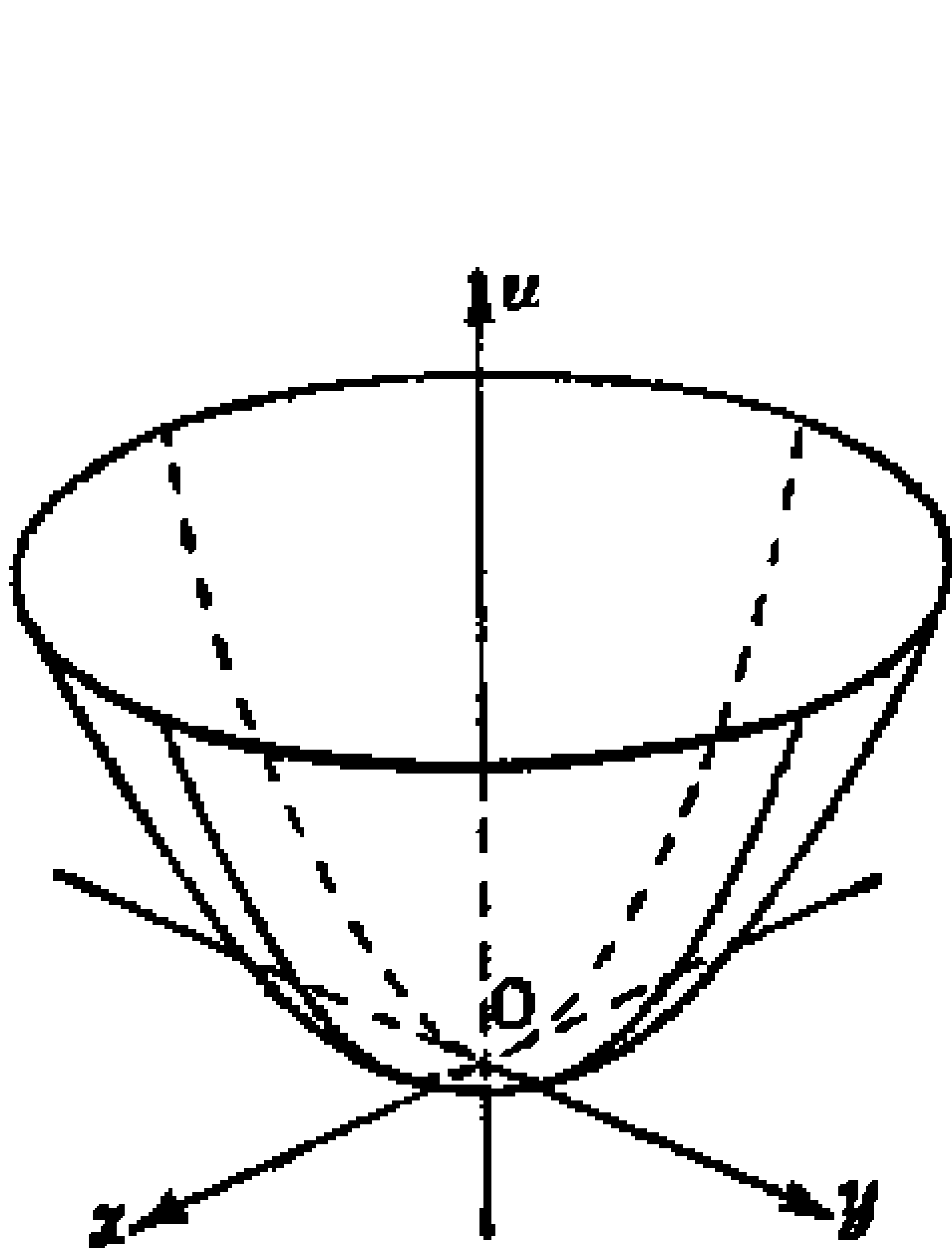


图 1.9　$u=x^2+y^2$

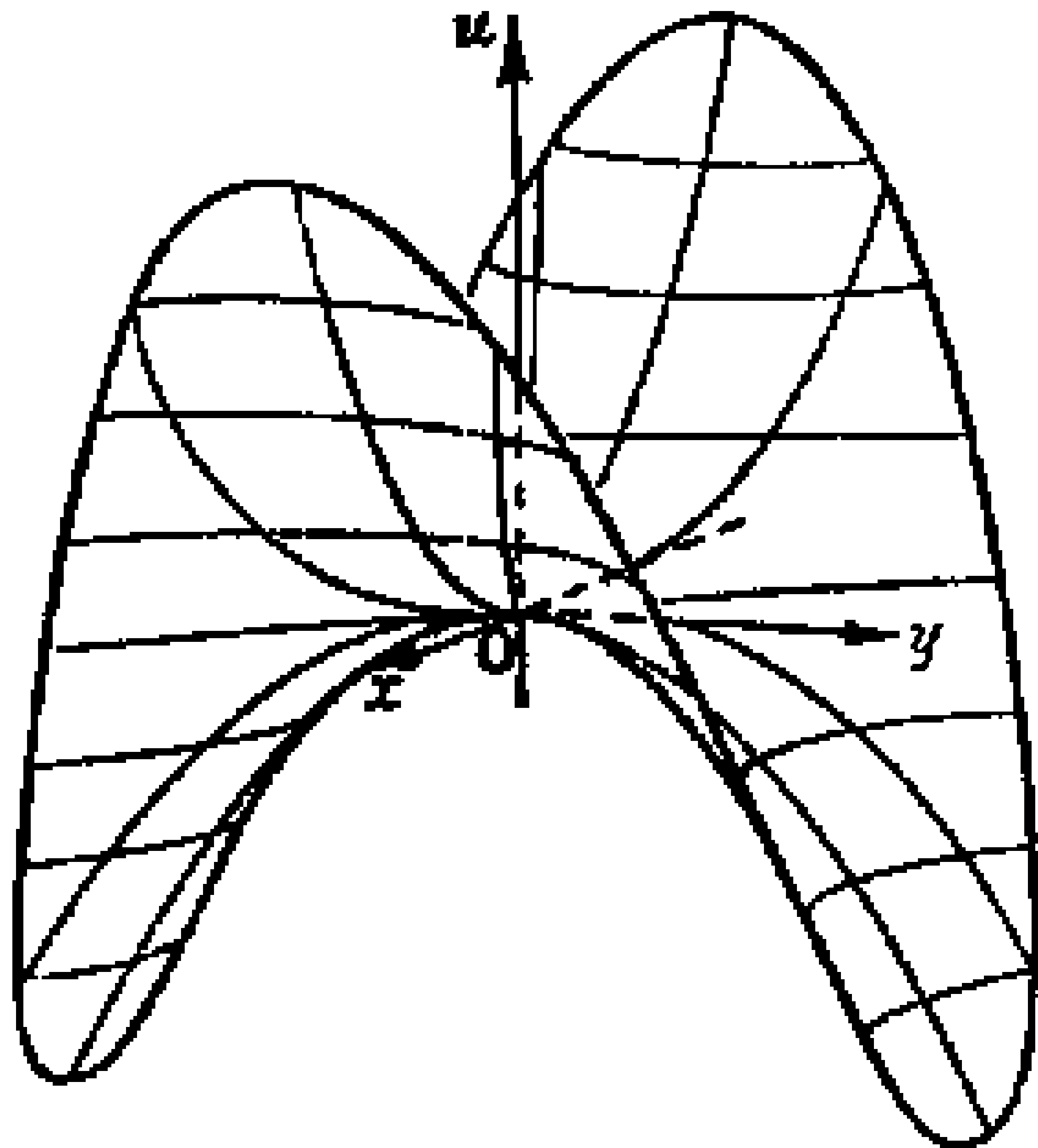


图 1.10　$u=x^2-y^2$

然而这种由直角坐标表示的方法有两个缺点．首先，当我们处理三个或更多个自变量的情形时，就不能用几何直观法．其次，即使是两个自变量，也常常把讨论局限于 xy 平面更为方便，因为在平面上我们进行几何构图不存在困难．从这一观点出发，二元函数

有些时候用一种等值线的几何表示法更好一些．在 xy 平面上，我们取所有那些使 $u=f(x,y)$ 为一常数值的那些点，如 $u=k$．这些点通常位于一条或几条曲线上，这就叫做函数的给定常数值 k 的**等值线** 或 **等高线**．我们也可以用平行于 xy 平面的平面 $u=k$ 来切割曲面 $u=f(x,y)$ 得到曲线，并把相交曲线垂直投射到 xy 平面来获得这些曲线．

对应高度 k 的值为 $k_1,k_2,\cdots$ 的这一族等高线，给我们提供了一种函数的表示法．在实践中 k 被指定一系列算术级数值，如 $k=\nu h$，其中 $\nu=1,2,\cdots$．于是等高线之间的距离给我们一个关于曲面 $u=f(x,y)$ 的陡度的度量，每两条相邻线之间，函数改变的值都相同．等高线靠得很近的地方，说明函数上升或下降很陡；等高线分得很开的地方，说明曲面是平坦的．这便是制作等高线地图(例如美国地质测量局的等高线地图) 的原理．

按照这个方法，线性函数 $u=ax+by+c$ 就被表示成一族平行直线 $ax+by+c=k$．函数 $u=x^2+y^2$ 就被表示成一族同心圆(图 1.11)，函数 $u=x^2-y^2$，其曲面为“马鞍形”(图 1.10)，就被表示成一族双曲线，如图 1.12 所示．

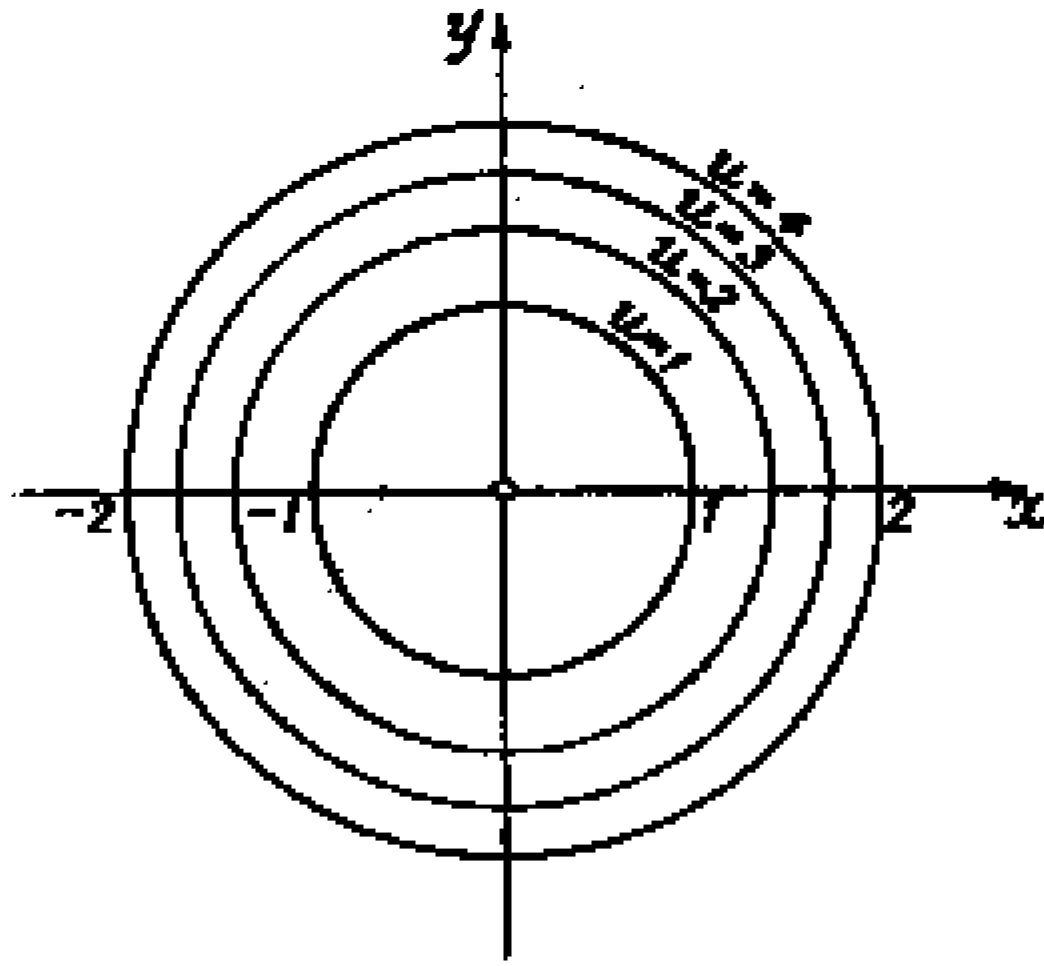


图 1.11　$u=x^2+y^2$ 的等高线

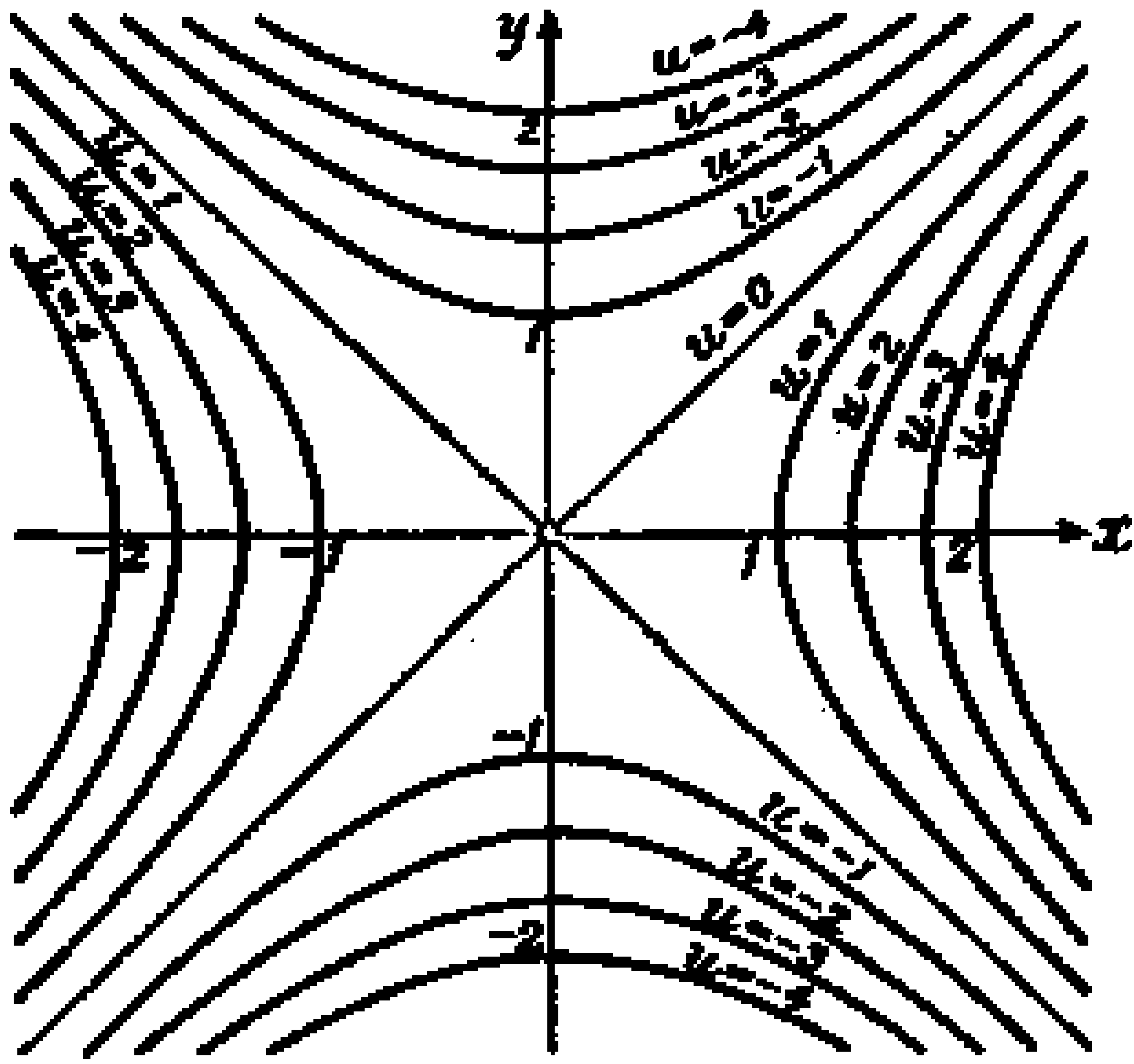


图 1.12　$u=x^2-y^2$ 的等高线

用等高线来表示函数 $u=f(x,y)$ 的方法，具有能推广到三个自变量的函数上去的优点．这里代替等高线的有 **等值曲面** $f(x,y,z)=$

k, 其中 k 为常数, 我们可以对它指定任何一系列适当的值. 例如, 函数 $u = x^2 + y^2 + z^2$ 的等值曲面是 (x, y, z) 坐标系中, 中心在原点的同心球.

练　习　1.2

1. 求下列函数在指定点处的值:

(a) $z = \left(\dfrac{\operatorname{arc\,cot}(x+y)}{\operatorname{arc\,cot}(x-y)}\right)^2$, 当 $x = \dfrac{1+\sqrt{3}}{2}$, $y = \dfrac{1-\sqrt{3}}{2}$;

(b) $w = e^{\cos z(x+y)}$, 当 $x = y = \dfrac{\pi}{2}$, $z = -1$;

(c) $z = y^{x \cos x^y}$, $x = e$, $y = \log \pi$;

(d) $z = \cos h(x+y)$, $x = \log \pi$, $y = \log \dfrac{1}{2}$;

(e) $z = \dfrac{x+y}{x-y}$, $x = \dfrac{1}{2}$, $y = \dfrac{1}{3}$.

2. 像第一卷中一样, 除了我们明确声明外, 用公式法确定的函数的定义域是指那些使表示式有意义的全部点的集合. 指出下列函数的定义域和值域:

(a) $z = \sqrt{x+y}$,

(b) $z = \sqrt{2x - y^2}$,

(c) $z = \dfrac{1}{\sqrt{x+y}}$,

(d) $z = \sqrt{1 - \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2}}$,

(e) $z = \log(x + 5y)$,

(f) $z = \sqrt{x \sin y}$,

(g) $w = \sqrt{a^2 - x^2 - y^2 - z^2}$,

(h) $z = \dfrac{x^2 - y^2}{x+y}$,

(i) $z = \sqrt{3 - x^2 - 2y^2}$,

(j) $z = \sqrt{-x^2 - y^2}$,

(k) $z = \log(x^2 - y^2)$,

(l) $z = \arctan \dfrac{x^2}{x^2 + y^2}$,

(m) $z = \arctan \dfrac{x}{x+y}$,

(n) $z = \cos \arctan \dfrac{y}{x}$,

(o) $z = \operatorname{arc\,cos} \log(x+y)$,

(p) $z = \sqrt{y \cos x}$.

3. 两个变量的一般的 n 次多项式的系数的个数是多少? 三个变量的情形呢? k 个变量的情形呢?

4. 对于下列函数, 画出对应于 $Z = -2, -1, 0, 1, 2, 3$ 的等高线的略图:

(a) $z=x^2y$,

(b) $z=x^2+y^2-1$,

(c) $z=x^2-y^2$,

(d) $z=y^2$,

(e) $z=y\left(1-\dfrac{1}{x^2+y^2}\right)$.

5. 画出函数 $z=\cos(2x+y)$ 对应于 $z=0,\pm1,\pm\dfrac{1}{2}$ 的等高线.

6. 画出下列曲面的略图:

(a) $z=2xy$,

(b) $z=x^2+y^2$,

(c) $z=x-y$,

(d) $z=x^2$,

(e) $z=\sin(x+y)$.

7. 求函数

$$z=\log\frac{1+\sqrt{x^2+y^2}}{1-\sqrt{x^2+y^2}}$$

的等值线.

8. 求使得函数 $u=2(x^2+y^2)/Z$ 等于常数的曲面.

1.3 连 续 性

a. 定义

像在一元函数的理论中一样，当我们考虑多元函数时，连续性的概念也是突出的. 说函数 $u=f(x,y)$ 在点 (ξ,η) 处连续，粗略地说，意味着在所有靠近 (ξ,η) 的点 (x,y) 处， $f(x,y)$ 的值与 $f(\xi,\eta)$ 相差很小. 我们把这个概念更精确地表述如下: 设 f 的定义域为 R, 且 $Q=(\xi,\eta)$ 是 R 中的一个点, 如果对每一个 $\varepsilon>0$ 都存在一

个 $\delta > 0$, 使得对 R 中所有的点 $P(x,y)$ 只要 [1]

$$\overline{PQ} = \sqrt{(x-\xi)^2 + (y-\eta)^2} < \delta \tag{1}$$

就有

$$|f(P) - f(Q)| = |f(x,y) - f(\xi,\eta)| < \varepsilon \tag{2}$$

成立, 则 f 在点 Q 连续.

若一个函数在一个点集 D 中的每一个点处都连续, 我们就说它在 D 中连续.

下面的事实几乎是明显的: 即连续函数的和、差与积也都是连续的. 连续函数的商, 在那些分母不为零的点处定义一个连续函数 (证明见下一节第 23 页). 特别是, 所有多项式都连续, 所有有理函数在分母不为零的那些点处都连续. 还有, 连续函数的连续函数本身是连续的 (参阅第 23 页).

一个多元函数的不连续点的类型可能比一元函数的复杂得多. 举例来说, 不连续点可能出现在整个曲线弧上而不仅仅是在一些孤立点上. 这种情形发生在由下式定义的函数:

$$\begin{cases} u = y/x, & \text{当 } x \neq 0, \\ u = 0, & \text{当 } x = 0. \end{cases}$$

它沿着整个直线 $x = 0$ 不连续. 更有甚者, 一个函数 $f(x,y)$ 可能对每一个固定的 y 值, 关于 x 连续, 且对每一个固定的 x 值, 关于 y 连续, 但作为点 (x,y) 的一个二元函数仍然是不连续的. 举例来说

$$f(x,y) = \frac{2xy}{x^2+y^2} \text{ 当 } (x,y) \neq (0,0),$$
$$f(0,0) = 0.$$

---

1) 我们可以用小正方形代替限制 (x,y) 在以 (ξ,η) 为中心的小圆盘. 连续性定义中的条件 (1) 就可改为

$$|x-\xi| < \delta \text{ 且 } |y-\eta| < \delta. \tag{1'}$$

这个函数就任何固定的 $y \neq 0$ 而言，显然关于 x 是连续的，因为分母不会为 0. 对 $y=0$, 我们可得 $f(x,0)=0$, 它依然是一个关于 x 的连续函数. 同理，$f(x,y)$ 对任何固定的 x 是一个关于 y 的连续函数. 但是在直线 $y=x$ 上，除了 $x=y=0$ 这点以外所有的点，我们都有 $f(x,y)=1$, 并且在这条直线上有任意靠近原点的点，因此 $f(x,y)$ 在点 $(0,0)$ 处不连续.

如同一元函数的情形一样，我们说一个函数 $f(P)=f(x,y)$ 在 xy 平面的集合 R 内 **一致连续**, 是指：如果 f 在 R 的全部点上有定义，并且对每一个 $\varepsilon>0$, 存在一个正数 $\delta=\delta(\varepsilon)$, 使得对 R 内的任何两个点 P 与 Q, 其距离 $<\delta$ 时，都有 $|f(P)-f(Q)|<\varepsilon$. [1] 量 $\delta=\delta(\varepsilon)$ 叫做 f 的一个 **连续模**. 我们有如下的基本定理：

若一个函数 f 在有界闭集合 R 上有定义并且连续，则在 R 内一致连续(其证明可参阅本章附录).

特别重要的情形是，我们可以找到一个与 ε 成比例的连续模(见第一卷第 45 页). 如果函数 $f(P)$ 在 R 中有定义，并且存在一个常量 L 使得

$$|f(P)-f(Q)| \leq L\overline{PQ} \tag{3}$$

对 R 内的一切点 P,Q 成立，则称这函数为 **利普希茨连续** 函数. (L 叫做"利普希茨常数", 关系式 (3) 叫做"利普希茨条件".) 很清楚，每一个利普希茨连续函数 f 是一致连续的，并且有连续模 $\delta=\varepsilon/L$. [2]

---

1) 一致连续的要点是 δ 依赖于 ε 而不依赖于 P 或 Q.

2) 当我们用霍尔德尔条件代替利普希茨条件 (3) 时，可以得到更广泛的一类"霍尔德尔连续"函数 f. 霍尔德尔条件是指：

$$|f(P)-f(Q)| \leq L\overline{PQ}^{\alpha},\quad P,Q \text{ 在 } R \text{ 内},$$

这里 L 和 α 是常数，且 $0<\alpha\leq 1$ (参看第一卷第 46 页), 这些函数也是一致连续的，并且我们可以取连续模为

$$\delta=(\varepsilon/L)^{1/\alpha}.$$

b. 多元函数的极限概念

函数的极限概念与其连续性概念密切相关. 我们假设, 函数 $f(x,y)$ 的定义域为 $R, Q=(\xi,\eta)$ 为 R 的闭包上的一个点. 如果对于每一个 $\varepsilon>0$, 我们可以找到 (ξ,η) 的一个邻域

$$PQ=\sqrt{(x-\xi)^2+(y-\eta)^2}<\delta, \tag{4}$$

使对于所有属于 R 且在这个邻域内的点 $P=(x,y)$, [1] 都有

$$|f(P)-L|=|f(x,y)-L|<\varepsilon,$$

我们就说, **当 (x,y) 趋近于 (ξ,η) 时 f 有极限 L, 并写成** [2]

$$\lim_{(x,y)\to(\xi,\eta)} f(x,y)=L \text{ 或 } \lim_{P\to Q} f(P)=L. \tag{5}$$

在点 (ξ,η) 属于 f 的定义域的情形, $(x,y)=(\xi,\eta)$ 是属于 R 的一个点, 对于一切 $\delta>0$ 都满足 (4). 这时 (5) 特别隐含着

$$|f(\xi,\eta)-L|<\varepsilon$$

对所有的 $\varepsilon>0$ 都成立, 因此 $L=f(\xi,\eta)$. 然而根据定义, 关系式

$$\lim_{(x,y)\to(\xi,\eta)} f(x,y)=f(\xi,\eta)$$

与 f 在点 (ξ,η) 连续的条件相同. 因此, **函数 f 在点 (ξ,η) 处连续等价于 f 在 (ξ,η) 有定义并且当 (x,y) 趋近于 (ξ,η) 时 $f(x,y)$ 有极限 $f(\xi,\eta)$.**

若 f 在其定义域的边界点 (ξ,η) 上没有定义, 但当 $(x,y)\to(\xi,\eta)$ 时有一个极限 L, 我们就自然可以令 $f(\xi,\eta)=L$ 而将 f 的

---

1) 对于区域 R 的**外点** (ξ,η), 这个概念是没有意义的, 因为不存在那些使 f 有定义的并可任意接近 (ξ,η) 的点, 因而每一个 L 都可以认为是极限.

2) 或者另外记作 $\lim f(x,y)=L$, 当 $(x,y)\to(\xi,\eta)$, 或者记作 $\lim\limits_{\substack{x\to\xi\\ y\to\eta}} f(x,y)=L$.

定义加以扩展；函数 f 按这个方式扩展之后，就在 (ξ,η) 处连续. 若 $f(x,y)$ 在其定义域 R 中连续，我们可以用极限值来扩展 f 的定义，不仅对 R 的一个边界点 (ξ,η)，而是同时对 R 的所有能使 f 有极限的边界点. 扩展以后的函数仍然是连续的，读者可以作为一个练习来证明. 举一个例来说，函数

$$f(x,y)=e^{-x^2/y}$$

对 $y>0$ 的所有 (x,y) 有定义. 这个函数显然在其定义域 R (即上半平面) 的所有点连续. 考虑一个边界点 $(\xi,0)$. 很清楚，当 $\xi\neq 0$ 时我们有

$$\lim_{(x,y)\to(\xi,0)} f(x,y)=\lim_{s\to+\infty} e^{-s}=0,$$

这里 y 限于取正值. 如果我们这样来定义扩展后的函数 $f^*(x,y)$，即对 $y>0$ 和所有的 x 令

$$f^*(x,y)=f(x,y)=e^{-x^2/y},$$

而对 $x\neq 0$ 令

$$f^*(x,0)=0,$$

那么函数 f^* 在其定义域 R^* 内连续，这里 R^* 为封闭上半平面 $y\geq 0$，除去点 $(0,0)$. 在原点处， f^* 没有极限，因此不能采用这种方式来定义 $f^*(0,0)$，使扩展后在原点连续. 事实上，对抛物线 $y=kx^2$ 上的点 (x,y)，我们有

$$f(x,y)=e^{-1/k},$$

它沿着不同的抛物线接近原点时有不同的极限值，所以当 $(x,y)\to(0,0)$ 时， $f(x,y)$ 不存在唯一的极限.

我们还可以把**函数** $f(x,y)$ **的极限**概念联系到**序列的极限**(参阅第一卷第 89 页). 设 f 的定义域为 R，且

$$\lim_{(x,y)\to(\xi,\eta)} f(x,y)=L.$$

设 $P_n = (x_n, y_n)\ (n = 1, 2, \cdots)$ 为 R 内任何一个点序列，且 $\lim\limits_{n\to\infty} P_n = (\xi, \eta)$，则数序列 $f(x_n, y_n)$ 有极限 L. 这是因为对所有 R 内足够靠近 (ξ, η) 的点 (x, y)，$f(x, y)$ 与 L 的差可以任意小，而只要 n 足够大，(x_n, y_n) 就足够靠近 (ξ, η). 反之，如果对于 R 内的每一个以 (ξ, η) 为极限的点序列 (x_n, y_n)，我们都有

$$\lim_{n\to\infty} f(x_n, y_n) = L,$$

那么当 $(x, y) \to (\xi, \eta)$ 时，$\lim\limits_{(x,y)\to(\xi,\eta)} f(x, y)$ 存在且其值为 L. 读者可以容易地作出证明. 如果我们限制点 (ξ, η) 在 f 的定义域 R 内，我们就得到以下论断：f **在它的定义域** R **中连续恰好意味着，当** $\lim(x_n, y_n) = (\xi, \eta)$ **时**

$$\lim_{n\to\infty} f(x_n, y_n) = f(\xi, \eta), \tag{6}$$

或者说

$$\lim_{n\to\infty} f(x_n, y_n) = f\left(\lim_{n\to\infty} x_n, \lim_{n\to\infty} y_n\right),$$

这里我们只考虑 R 中的这种收敛序列 (x_n, y_n)，并且它们的极限也在 R 中. 于是，从本质上来说，函数 f 连续就允许把 f 的记号和极限的记号互换次序.

很清楚，当函数 f 的定义域不是一个二维区域而是一条曲线或任何其他点集的时候，函数极限和连续的概念也适用. 例如，函数

$$f(x + y) = (x + y)!$$

在包括所有直线 $x + y = 常数 = n$（n 为正整数）的集合 R 内有定义；很显然，f 在其定义域 R 中连续.

前面（见第 19 页）曾叙述过，当 $f(x, y)$ 和 $g(x, y)$ 在一点 (ξ, η) 处连续时，则 $f + g, f - g, f \cdot g$ 以及当 $g(\xi, \eta) \neq 0$ 时的 f/g 都在 (ξ, η) 处连续. 这些法则可以直接从收敛序列所表述的连续性推导

出来. 对于任何属于 f 和 g 的定义域, 且收敛到 (ξ,η) 的点序列 (x_n,y_n), 根据公式 (6) 我们有

$$\lim_{n\to\infty} f(x_n,y_n) = f(\xi,\eta), \quad \lim_{n\to\infty} g(x_n,y_n) = g(\xi,\eta);$$

从而 $f(x_n,y_n)+g(x_n,y_n)$ 等等的收敛性可以从序列的运算法则推导出来 (第一卷第 77 页).

c. 无穷小函数的阶

若函数 $f(x,y)$ 在点 (ξ,η) 处连续, 则当 x 趋于 ξ 且 y 趋于 η 时, 差 $f(x,y)-f(\xi,\eta)$ 趋于 0. 引入新的变量 $h=x-\xi$ 和 $k=y-\eta$, 我们就可以把这表示如下: 当 h 和 k 趋向于 0 时, 变量 h 和 k 的函数 $\phi(h,k)=f(\xi+h,\eta+k)-f(\xi,\eta)$ 趋于 0.

我们将常常遇到这样的函数 $\phi(h,k)$, 当 h 和 k 趋于 0 时它趋于 0. 像在一元的情况一样, 为了各种目的, 经常把 $\phi(h,k)$ 当 $h\to 0$ 和 $k\to 0$ 时的趋于 0 的情况更精确地用 "无穷小量的阶" 或 $\phi(h,k)$ 的 "模量的阶" 来加以区别. 为此, 我们以坐标为 $x=\xi+h$ 和 $y=\eta+k$ 的点到坐标为 $x=\xi$ 和 $y=\eta$ 的点的距离

$$\rho=\sqrt{h^2+k^2}=\sqrt{(x-\xi)^2+(y-\eta)^2}$$

作为比较的基础, 并采用下述定义:

当 $\rho\to 0$ 时函数 $\phi(h,k)$ 是无穷小, 只要存在一个独立于 h 和 k 的常数 C, 使得不等式

$$\left|\frac{\phi(h,k)}{\rho}\right|\le C$$

对于足够小的值 ρ 都成立, 那么它与 $\rho=\sqrt{h^2+k^2}$ 至少有相同的阶; 也就是说, 有一个 $\delta>0$ 使对于所有适合 $0<\sqrt{h^2+k^2}<\delta$ 的 h 和 k 的值, 上述不等式都成立. 我们用记号写作: $\phi(h,k)=O(\rho)$. 进一步, 如果当 $\rho\to 0$ 时, 商式 $\phi(h,k)/\rho$ 趋于 0, 我们就说 $\phi(h,k)$

是比 ρ 高阶的无穷小[1]. 这将用记号表示成

$$\phi(h,k)=o(\rho)$$

当 $(h,k)\to 0$ (见第一卷第 283 页, 那里记号 "o" 和 "O" 是对一元函数作解释的).

让我们来研究几个例子. 因为

$$\frac{|h|}{\sqrt{h^2+k^2}}\leq 1 \text{ 和 } \frac{|k|}{\sqrt{h^2+k^2}}\leq 1,$$

距离 ρ 在 x 轴和 y 轴方向上的分量 h 和 k 至少与距离本身具有相同的阶. 同样, 一个以 a、 b 为常系数的线性齐次函数 $ah+bk$ 或函数 $\rho\sin 1/\rho$ 也至少与距离本身有相同的阶. 对于大于 1 的固定值 a, 距离的幂 ρ^a 是比 ρ 高阶的无穷小; 用记号表示成: $\rho^a=o(\rho)$ 当 $a>1$. 类似地, 变量 h 和 k 的二次齐次多项式 $ah^2+bhk+ck^2$ 当 $\rho\to 0$ 时是一个比 ρ 高阶的无穷小:

$$ah^2+bhk+ck^2=o(\rho).$$

更一般化, 可用下述定义. 设在原点处一个足够小的圆内, 对所有非零的 (h,k) 值函数 $\omega(h,k)$ 有定义并且不等于 0. 如果对于某个合适选择的常数 C, 在点 $(h,k)=(0,0)$ 的邻域内, 关系式

$$\left|\frac{\phi(h,k)}{\omega(h,k)}\right|\leq C$$

成立, 则 $\phi(h,k)$ 当 $\rho\to 0$ 时, 至少与 $\omega(h,k)$ 同阶. 我们把这表示成符号方程 $\phi(h,k)=O(\omega(h,k))$. 同样, 若当 $\rho\to 0$ 时 $\dfrac{\phi(h,k)}{\omega(h,k)}\to 0$, 则 $\phi(h,k)$ 是比 $\omega(h,k)$ 高阶的无穷小, 记作

$$\phi(h,k)=o(\omega(h,k)).$$

---

1) 为了避免混淆, 我们明白地指出, 当 $\rho\to 0$ 时更高阶无穷小意味着在 $\rho=0$ 的邻域中有较小的值; 例如 ρ^2 是比 ρ 高阶的无穷小, 当 ρ 接近于 0 时, ρ^2 小于 ρ.

例如齐次多项式 $ah^2+bhk+ck^2$ 至少与 ρ^2 同阶，因为

$$|ah^2+bhk+ck^2| \le \left(|a|+\frac{1}{2}|b|+|c|\right)(h^2+k^2).$$

同样，$\rho = o(1/|\log\rho|)$，因为 $\lim\limits_{\rho\to 0}(\rho\log\rho)=0$ (第一卷第 283 页).

练 习 1.3

1. 函数 $z=(x-y)/(x+y)$ 沿 $y=-x$ 是不连续的. 描出该曲面当 $z=0,\pm1,\pm2$ 时的等值线. 当 $z=\pm m$ 而 m 很大时，等值线是什么样的曲线？

2. 考察函数 $z=(x^2+y)-\sqrt{x^2+y^2}$ 的连续性，这里当 $x=y=0$ 时 $z=0$. 描出 $z=k$ $(k=-4,-2,0,2,4)$ 时的等值线. 当 $y=-2,-1,0,1,2$ 时，z 只是 x 的函数，展示这函数的性态 (在同一张图纸上). 同样地，对于 $x=0,\pm1,\pm2$，展示 z 作为只是 y 的一个函数的图形. 最后，描出当 θ 是常数时 z 作为只是 ρ 的一个函数的图形 (ρ,θ 是极坐标).

3. 用确定连续模 $\delta(\varepsilon)$ 的方法验证下列函数在原点处是连续的：

(a) $f(x,y)=x^3-3xy^2$,

(b) $g(x,y)=x^4-6x^2y^2+y^4$.

指出每一个函数在原点是几阶无穷小.

4. 证明下列函数是连续的：

(a) $\sin(x^2+y)$, (b) $\dfrac{\sin xy}{\sqrt{x^2+y^2}}$,

(c) $\dfrac{x^3+y^3}{x^2+y^2}$, (d) $x^2\log(x^2+y^2)$.

对于每一种情形，在 (0,0) 点我们定义函数值等于给定表达式的极限.

5. 对于下列连续函数，求出连续模 $\delta=\delta(\varepsilon,x,y)$ 的值：

(a) $f(x,y)=\sqrt{1+x^2+2y^2}$.

(b) $f(x,y)=\sqrt{1+e^{xy}}$.

6. 函数 $z=1/(x^2-y^2)$ 在哪儿是不连续的？

7. 函数 $z=\tan\pi y/\cos\pi x$ 在哪儿是不连续的？

8. 求函数 $z=\sqrt{y\cos x}$ 的连续点 (x,y) 的集合.

9. 证明函数 $z=1/(1-x^2-y^2)$ 在圆盘 $x^2+y^2<1$ 内是连续的.

10. 求多项式

$$P=ax^2+2bxy+cy^2$$

在 $x=0, y=0$ 的邻域内与 ρ^2 同阶的条件 (即 P/ρ^2 与 ρ^2/P 是有界的).

11. 确定下列函数是否是连续的；如果不连续，它们在哪儿是不连续的:

(a) $\sin\dfrac{y}{x}$, (b) $\dfrac{x^3+y^2}{x^2+y^2}$,

(c) $\dfrac{x^3+y^2}{x^3+y^3}$, (d) $\dfrac{x^3+y^2}{x^2+y}$.

12. 证明当 (x,y) 沿任何一条直线趋于原点时，函数

$$f(x,y)=\frac{x^4y^4}{(x^2+y^4)^3},\quad g(x,y)=\frac{x^2}{x^2+y^2-x}$$

趋于 0; 但 f 和 g 在原点是不连续的.

13. 判定下列函数在 $x=y=0$ 点是否有极限；如果极限存在，求出该极限.

(a) $\dfrac{x^2-y^2}{x^2+y^2}$, (b) $\dfrac{x^2+2xy+y^2}{x^2+y^2}$,

(c) $\dfrac{x^2+3xy+y^2}{x^2+4xy+y^2}$, (d) $-\dfrac{|x-y|}{x^2-2xy+y^2}$,

(e) $\exp[-|x-y|/(x^2-2xy+y^2)]$,

(f) $|x|^y$, (g) $|x|^{|1/y|}$,

(h)* $\dfrac{|y|^{|x|}\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}+|y/x|}$.

14. 对于习题 13 中的那些在 $x=y=0$ 点处有极限的函数，求出它们的连续模 $\delta(\varepsilon)$，这些函数在原点的值定义为他们的极限值.

15. 证明函数 $f(x,y,z)=(x^2+y^2-z^2)/(x^2+y^2+z^2)$ 在点 $(0,0,0)$ 处是不连续的.

16. 若 $P(x,y)$ 和 $Q(x,y)$ 都是 n $(n>0)$ 次多项式，在原点其值为 0，证明

$$R(x,y)=\frac{P(x,y)}{Q(x,y)}$$

在原点是不连续的.

17. 当 (x,y) 以任何方式趋于 $(0,0)$ 时，求下列表达式的极限：

(a) $\dfrac{\sin(x^2+y^2)}{x^2+y^2}$，(b) $\dfrac{\sin(x^4+y^4)}{x^2+y^2}$，(c) $\dfrac{e^{-1/(x^2+y^2)}}{x^4+y^4}$.

18. 证明当 (x,y) 趋于 $(0,0)$ 时，函数 $z=3(x-y)/(x+y)$ 可趋于任何一个极限. 给出 (x,y) 的变化情况使分别适合下列条件：

(a) $\lim\limits_{\substack{x\to0\\y\to0}} z=2$，(b) $\lim\limits_{\substack{x\to0\\y\to0}} z=-1$，(c) $\lim\limits_{\substack{x\to0\\y\to0}} z$ 不存在.

19. 如果当 $(x,y)\to(0,0)$，沿每一条经过原点的直线都有 $f(x,y)\to0$，是否当 $(x,y)\to(0,0)$ 沿任何一条路径都有 $f(x,y)\to 0$.

20. 在原点 $(0,0)$ 的邻域内考察 $z=y\log x$ 的性质.

21. 设 $z=f(x,y)=(x^2-y)/2x$，画出下列各式的图形：

(a) $z=f(x,x^2)$，　(b) $z=f(x,0)$，

(c) $z=f(x,1)$，　(d) $z=f(x,x)$.

当 $(x,y)\to(0,0)$ 时，$f(x,y)$ 的极限是否存在？

22. 给出下述情况的几何解释：$\phi(h,k)$ 与 $\rho=\sqrt{h^2+k^2}$ 是同阶无穷小.

问　题　1.3

1. 让我们把连续函数 f 这样开拓到函数 f^*：在 f 的定义域内

定义 $f^* = f$，在边界点上定义 $f^*(Q) = \lim\limits_{P \to Q} f(P)$，只要这极限存在. 证明 f^* 是连续的.

2. 证明，对于 $(x, y) \to (\xi, \eta)$，极限 $\lim f(x, y)$ 存在且其值为 L，当且仅当对于 f 的定义域中的每一个以 (ξ, η) 为极限的点序列 (x_n, y_n) 都有 $\lim\limits_{n \to \infty} f(x_n, y_n) = L$.

1.4 函数的偏导数

a. 定义. 几何表示

如果在多元函数中，除了一个自变量之外，我们指定所有其他自变量以确定的值，而只允许那一个自变量，譬如 x 变动，这函数就成为一个一元函数. 我们来研究两个变量 x 和 y 的一个函数 $u = f(x, y)$，并指定 y 以一个确定的固定值 $y = y_0 = c$. 得出的函数 $u = f(x, y_0)$ 为单独一个变量 x 的函数，它可以几何地说成是用平面 $y = y_0$ 去切割曲面 $u = f(x, y)$ (参阅图 1.13 和 1.14). 这样形成的平面上的交线可以用方程 $u = f(x, y_0)$ 表示. 如果我们对这个方程用普通方法在点 $x = x_0$ 上求导，假设 f 在 (x_0, y_0) 的邻域有定义，且导数存在 [1]，我们就得到 $f(x, y)$ 在点 (x_0, y_0) 处**关于 x 的偏导数**

$$\lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}.$$

几何上看，这个偏导数表示 x 轴的一条平行线与曲线 $u = f(x, y_0)$ 的切线之间的夹角的正切. 所以它是**曲面** $u = f(x, y)$ **在 x 轴方向的斜率**.

为了表示这些偏导数，常用几种不同的符号，其中一种是

$$\lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = f_x(x_0, y_0) = u_x(x_0, y_0).$$

---

1) 我们将不试图去定义区域的边界点上的导数 (除了偶尔看作是内点逼近于边界点时偏导数的极限).

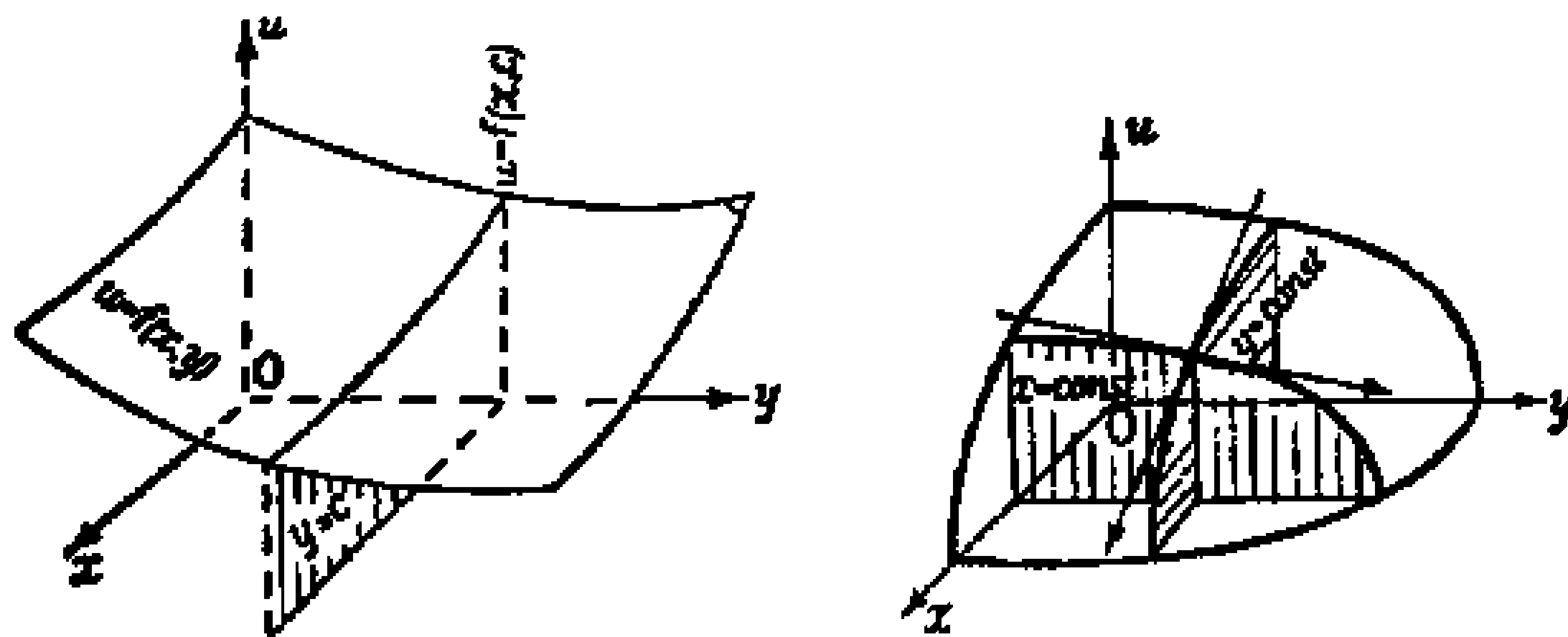


图 1.13 和图 1.14　$u=f(x,y)$ 的横断面

如果我们希望强调偏导数是一种差商的极限，我们可以把它表示成

$$\frac{\partial f}{\partial x} \text{ 或 } \frac{\partial}{\partial x}f.$$

这里我们用一个特别的圆形字母 ∂ 以代替在一元函数微分法中常用的 d，这是为了表示我们处理的是多元函数，而关于其中一个求导.

为了某些目的，可方便地应用柯西符号 D (已在第一卷第 179 页上提到过)，并记作

$$\frac{\partial f}{\partial x}=D_x f.$$

但我们很少用这个符号.

完全相同的方式，我们用下面的关系式来定义 $f(x,y)$ 在点 (x_0,y_0) 处对 y 的偏导数：

$$\lim_{h\to 0}\frac{f(x_0,y_0+h)-f(x_0,y_0)}{h}=f_y(x_0,y_0)=D_y f(x_0,y_0).$$

这表示曲面 $u=f(x,y)$ 与垂直于 x 轴的平面 $x=x_0$ 的交线的斜率 (图 1.14).

现在让我们把迄今为止认为是固定的点 (x_0,y_0) 看作变量，因而把下标 0 省略掉. 换句话说，我们把求导运算看作是在 $f(x,y)$ 的定义域内任何一个点 (x,y) 处进行的. 于是两个导数

$$u_x(x,y)=f_x(x,y)=\frac{\partial f(x,y)}{\partial x},$$

$$u_y(x,y) = f_y(x,y) = \frac{\partial f(x,y)}{\partial y}$$

本身都是 (x,y) 的函数.

举例来说，函数 $u = x^2 + y^2$ 具有偏导数 $u_x = 2x$ (对 x 求导时，把 y^2 当作常数，因而导数为 0) 和 $u_y = 2y$. 函数 $u = x^3 y$ 的偏导数是 $u_x = 3x^2 y$ 和 $u_y = x^3$.

类似地，对于任何 n 个自变量的一个函数 $f(x_1, x_2, \cdots, x_n)$，我们定义它关于 x_1 的偏导数为

$$\begin{aligned}\frac{\partial f(x_1, x_2, \cdots, x_n)}{\partial x_1} &= \lim_{h\to 0} \frac{f(x_1+h, x_2, \cdots, x_n) - f(x_1, x_2, \cdots, x_n)}{h} \\ &= f_{x_1}(x_1, x_2, \cdots, x_n) = D_{x_1} f(x_1, x_2, \cdots, x_n),\end{aligned}$$

假设这极限存在.

当然我们还可以重复这样做，把"一阶"偏导数 $f_x(x,y)$ 和 $f_y(x,y)$ 再对一个自变量求导，得到 $f(x,y)$ 的**高阶偏导数**. 求导的顺序是用下标或分母中符号 ∂x 和 ∂y 从右到左的顺序来表示的[1]. 二阶偏导数用下列符号来表示:

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial^2 f}{\partial x^2} = f_{xx} = (D_x)^2 f, \\ \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial^2 f}{\partial x \partial y} = f_{xy} = D_x D_y f, \\ \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial^2 f}{\partial y \partial x} = f_{yx} = D_y D_x f, \\ \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial^2 f}{\partial y^2} = f_{yy} = (D_y)^2 f.\end{aligned}$$

三阶偏导数用下列符号表示:

$$\begin{aligned}\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x^2}\right) &= \frac{\partial^3 f}{\partial x^3} = f_{xxx}, \\ \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x^2}\right) &= \frac{\partial^3 f}{\partial y \partial x^2} = f_{yxx},\end{aligned}$$

---

1) 这同算子的符号乘积的一般记号是一致的 (参阅第一卷第 57 页). 其实求导的顺序在大多数我们感兴趣的情况下是不重要的 (参阅第 39 页).

$$\frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x \partial y}\right)=\frac{\partial^3 f}{\partial x^2 \partial y}=f_{xxy},$$

等等，一般地，n 阶偏导数用下列符号表示：

$$\frac{\partial}{\partial x}\left(\frac{\partial^{n-1} f}{\partial x^{n-1}}\right)=\frac{\partial^n f}{\partial x^n}=f_{x^n},$$

$$\frac{\partial}{\partial y}\left(\frac{\partial^{n-1} f}{\partial x^{n-1}}\right)=\frac{\partial^n f}{\partial y \partial x^{n-1}}=f_{yx^{n-1}},$$

等等.

偏导数的不同符号各有其优点. 把函数 $f(x,y)$ 对其第一个自变量的偏导数写成 $\frac{\partial f(x,y)}{\partial x}$ 或 $D_x f(x,y)$，强调了微分具有作用于函数的**算子** D_x 或 $\frac{\partial}{\partial x}$ 的性质，书写时作为**因子**乘到函数前面去. 高阶导数的记号保持了这种乘积的概念：

$$\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x} f\right)=\frac{\partial^2}{\partial y \partial x} f=D_y \cdot D_x f.$$

算子记号的缺点在于，当它要用来表示导数是对自变量的什么值而言的时候，它是笨拙的. 例如，若

$$f(x,y)=x^2+2xy+4y^2,$$

则其在点 $x=1, y=2$ 处对 x 的导数可写成

$$\left(\frac{\partial f(x,y)}{\partial x}\right)_{\substack{x=1\\y=2}}=f_x(1,2)=(2x+2y)_{\substack{x=1\\y=2}}=6.$$

我们不应当简单地写成

$$\frac{\partial f(1,2)}{\partial x},$$

因为 $f(1,2)$ 是常量 21，因而它对 x 的导数为 0. 正如一元的情形一样，一个函数具有导数是该函数的一个特殊性质，并不是所有的连续函数都具备的[1]. 然而，实际上重要的函数都具备这个性质，例外只发生在孤立的点或曲线上.

---

1) 这里我们不用“可微”一词，因为“可微”一词隐含着比关于 x 和关于 y 偏导数存在更多的东西，关于这一名词的解释请看第 43 页.

练 习 1.4 a

1. 求下列函数的偏导数 $\partial z/\partial x$, $\partial z/\partial y$:

(a) $z = ax^n + by^m$, a, b, m, n 是常数,

(b) $z = 2xe^{y^2} + 3y$,

(c) $z = 2\frac{x}{y} + 3\frac{y}{x}$,

(d) $z = \arctan\frac{y}{x^2}$,

(e) $z = x^2y^{3/2}$,

(f) $z = y^x$,

(g) $z = x^{1/2}y^{3/4}$,

(h) $z = 3^{x/y}$,

(i) $z = \log\left(x + \frac{y}{x^2}\right)$,

(j) $z = \cos(x^2 + y)$,

(k) $z = \tan(xy^3 + e^x)$,

(l) $z = \frac{\cos x}{\sin y}$,

(m) $z = xe^y + ye^x$,

(n) $z = x\sqrt{x^2 + y^2}$.

2. 求下列函数的一阶偏导数:

(a) $\sqrt[3]{x^2 + y^2}$, (b) $\sin(x^2 - y)$,

(c) e^{x-y}, (d) $\frac{1}{\sqrt{1 + x + y^2 + z^2}}$.

(e) $y\sin xz$, (f) $\log\sqrt{1 + x^2 + y^2}$.

3. 求下列函数所有的一、二阶偏导数:

(a) xy,

(b) $\log xy$,

(c) $\tan(\arctan x + \arctan y)$,

(d) x^y,

(e) $e^{(x^y)}$

4. 令 $w = f(x, y, z) = (\cos x/\sin y)e^z$. 求 f_x, f_y, f_z, 当 $x = \pi$, $y = \pi/2$, $z = \log 3$.

5. 设 $f(x, y) = y\cosh x + x\sinh y$, 在 $x = 0$, $y = 0$ 点求 $f_{x^2} + f_{y^2}$.

6. 证明函数 $u = e^x\cos y$, $v = e^x\sin y$ 满足条件 $u_x = v_y$, $u_y = -v_x$.

7. 证明练习 6 的函数满足偏微分方程

$$f_{xx} + f_{yy} = 0.$$

证明下列函数同样也满足上面这个方程:

(a) $\log\sqrt{x^2+y^2}$,

(b) $\arctan\dfrac{y}{x}$,

(c) $\dfrac{y}{x^2+y^2}$,

(d) $3x^2y-y^3$,

(e) $\sqrt{x+\sqrt{x^2+y^2}}$.

8. 令 $r=\sqrt{x^2+y^2+z^2}$, 求 $r_{xx}+r_{yy}+r_{zz}$.

9. 如果 $z=y^3+ayx^2$ 适合方程 $z_{xx}+z_{yy}=0$, 求常数 a.

10. 证明函数

$$f(x_1,x_2,\cdots,x_n)=\frac{1}{\left(x_1^2+x_2^2+\cdots+x_n^2\right)^{(n-2)/2}}$$

满足方程

$$f_{x_1x_1}+f_{x_2x_2}+\cdots+f_{x_nx_n}=0.$$

问　题　1.4 a

1. 一个三元函数有几个 n 阶偏导数? k 个变数的函数呢?

2. 举出这样的例子, 使函数 $f(x,y)$ 的 f_x 存在但 f_y 不存在.

3. 求函数 $f(x,y)$, 使它是 (x^2+y^2) 的函数并且也是乘积 $\psi(x)\psi(y)$ 的形式; 也就是求解未知函数方程

$$f(x,y)=\phi(x^2+y^2)=\psi(x)\psi(y).$$

4. 证明形如

$$u(x,y,z)=\frac{f(t+r)}{r}+\frac{g(t-r)}{r}$$

(其中 $r^2=x^2+y^2+z^2$) 的任一函数都满足方程

$$u_{xx}+u_{yy}+u_{zz}=u_{tt}.$$

b. 例

实际上，偏微分包含的内容读者在以前都已遇到过。因为，按照定义，除了我们要对它求导的一个自变量外，其他的自变量都保持常数。因此我们可以把其他自变数看作常数，而按照一元函数求导法则去求导。下面我们列出若干简单函数的一些偏导数。

1. 函数

$$f(x,y)=xy,$$

一阶（偏）导数

$$f_x=y, \quad f_y=x;$$

二阶（偏）导数

$$f_{xx}=0, \quad f_{xy}=f_{yx}=1, \quad f_{yy}=0.$$

2. 函数

$$f(x,y)=\sqrt{x^2+y^2},$$

一阶（偏）导数

$$f_x=\frac{x}{\sqrt{x^2+y^2}}, \quad f_y=\frac{y}{\sqrt{x^2+y^2}}.$$

[可见从原点到点 (x,y) 的向量半径 $r=\sqrt{x^2+y^2}$ 关于 x,y 的偏导数是 $\cos\phi=x/r$ 和 $\sin\phi=y/r$，这里 ϕ 是向量半径与 x 轴正方向的夹角.]

二阶（偏）导数

$$f_{xx}=\frac{y^2}{\sqrt{(x^2+y^2)^3}}=\frac{\sin^2\phi}{r},$$

$$f_{xy}=f_{yx}=-\frac{xy}{\sqrt{(x^2+y^2)^3}}=-\frac{\sin\phi\cos\phi}{r},$$

$$f_{yy}=\frac{x^2}{\sqrt{(x^2+y^2)^3}}=\frac{\cos^2\phi}{r}.$$

3. 三维空间中向量半径的倒数

$$f(x,y,z)=\frac{1}{\sqrt{x^2+y^2+z^2}}=\frac{1}{r},$$

一阶 (偏) 导数

$$f_x=-\frac{x}{\sqrt{(x^2+y^2+z^2)^3}}=-\frac{x}{r^3},$$

$$f_y=-\frac{y}{\sqrt{(x^2+y^2+z^2)^3}}=-\frac{y}{r^3},$$

$$f_z=-\frac{z}{\sqrt{(x^2+y^2+z^2)^3}}=-\frac{z}{r^3};$$

二阶 (偏) 导数

$$f_{xx}=-\frac{1}{r^3}+\frac{3x^2}{r^5},\quad f_{yy}=-\frac{1}{r^3}+\frac{3y^2}{r^5},\quad f_{zz}=-\frac{1}{r^3}+\frac{3z^2}{r^5},$$

$$f_{xy}=f_{yx}=\frac{3xy}{r^5},\quad f_{yz}=f_{zy}=\frac{3yz}{r^5},\quad f_{zx}=f_{xz}=\frac{3zx}{r^5}.$$

由此可见，函数

$$f=\frac{1}{\sqrt{x^2+y^2+z^2}}$$

除 (0,0,0) 外，对所有 (x,y,z) 都适合方程

$$f_{xx}+f_{yy}+f_{zz}=-\frac{3}{r^3}+\frac{3(x^2+y^2+z^2)}{r^5}=0;$$

我们说，函数 $f(x,y,z)=1/r$ 满足**偏微分方程**(“拉普拉斯 (Laplace) 方程”)

$$f_{xx}+f_{yy}+f_{zz}=0.$$

4. 函数

$$f(x,y)=\frac{1}{\sqrt{y}}e^{-(x-a)^2/4y},$$

一阶 (偏) 导数

$$f_x=\frac{-(x-a)}{2y^{3/2}}e^{-(x-a)^2/4y},$$

$$f_y = \left[-\frac{1}{2y^{3/2}} + \frac{(x-a)^2}{4y^{5/2}} \right] e^{-(x-a)^2/4y};$$

二阶 (偏) 导数

$$f_{xx} = \left(\frac{-1}{2y^{3/2}} + \frac{(x-a)^2}{4y^{5/2}} \right) e^{-(x-a)^2/4y},$$

$$f_{xy} = f_{yx} = \left(\frac{3(x-a)}{4y^{5/2}} - \frac{(x-a)^3}{8y^{7/2}} \right) e^{-(x-a)^2/4y},$$

$$f_{yy} = \left(\frac{3}{4} \frac{1}{y^{5/2}} - \frac{1}{2} \frac{(x-a)^2}{y^{7/2}} + \frac{(x-a)^4}{16y^{9/2}} \right) e^{-(x-a)^2/4y}.$$

因此偏微分方程

$$f_{xx} - f_y = 0$$

对于 x, y 恒成立.

c. 偏导数的连续性与存在性

对于一个一元函数来说，导数在某一点存在，就意味着函数在该点连续 (参阅第一卷第 187 页); 与此相反，偏导数的存在 **并不** 隐含二元函数连续. 例如函数

$$u(x,y) = \frac{2xy}{(x^2+y^2)}$$

且 $u(0,0) = 0$, 它处处都有偏导数，但我们已在第 19 页看到，它在原点是不连续的，从几何上来看，偏导数的存在只限制着函数在 x 轴方向和 y 轴方向上的性态，而不在任何其他方向上. 然而，如果具有 **有界** 偏导数，则确实隐含着函数连续，如以下定理所述:

若一个函数 $f(x,y)$ **在一个开集** R **内处处有偏导数** f_x **和** f_y, **并且各处的偏导数满足不等式**

$$|f_x(x,y)| < M, \quad |f_y(x,y)| < M,$$

其中 M 与 x 和 y 无关，则 $f(x,y)$ 在 R 内处处连续.[1)]

为了证明，我们考虑区域 R 内坐标分别为 (x,y) 和 $(x+h,y+k)$ 的两个点. 我们进一步假定，连接这两点到点 $(x+h,y)$ 的两线段完全在区域 R 内，这在点 (x,y) 是 R 的内点并且点 $(x+h,y+k)$ 与 (x,y) 足够靠近时肯定是这样. 于是我们有

$$\begin{aligned} &f(x+h,y+k)-f(x,y)\\ &\quad=\{f(x+h,y+k)-f(x+h,y)\}+\{f(x+h,y)-f(x,y)\}. \quad (7)\end{aligned}$$

右端第一个括号中的两项只是 y 不同；而第二个括号内的两项只是 x 不同. 因此，我们可以用通常的微分中值定理 (第一卷第 194 页)，于第一个括号中看作只是 y 的函数，于第二个括号中看作只是 x 的函数. 由此我们得出关系式

$$\begin{aligned} &f(x+h,y+k)-f(x,y)\\ &\quad=kf_y(x+h,y+\theta_1k)+hf_x(x+\theta_2h,y), \qquad (8)\end{aligned}$$

其中 θ_1,θ_2 为介于 0 到 1 之间的数. 换言之，对 y 的导数取连接 $(x+h,y)$ 到 $(x+h,y+k)$ 的垂直连线上一个点的值，而对 x 的导数取连接 (x,y) 到 $(x+h,y)$ 的水平连线上一个点的值. 根据假设，导数的绝对值均小于 M. 于是得出

$$|f(x+h,y+k)-f(x,y)|\le M(|h|+|k|). \qquad (9)$$

对于足够小的 h 和 k 的值，右边本身可以任意小，这就证明了 $f(x,y)$ 是连续的[2)].

---

1) 由这个定理的证明可以看出，这定理也适用于区域的边界点，只要能用一根由平行于坐标轴的两个线段组成的折线连接该边界点与区域内的任何一个邻近点，并且在该边界点上适当地定义 f.

2) 如果 f 的定义域是一个四边平行于坐标轴的矩形，这不等式就对定义域内任何两点 (x,y) 和 $(x+h,y+k)$ 都成立. 于是得出 f 还是利普希茨连续的 (见第 20 页).

练　习　1.4 c

1. 叙述并证明：三元函数 $f(x,y,z)$ 的一阶偏导数存在并有界，是 f 连续的一个充分条件.

2. 证明下列函数 $f(x,y)$ 连续：

(a) $f(x,y)=\begin{cases} e^{-1/(x^2+y^2)} & \text{当 } x,y\neq 0, \\ 0 & \text{当 } x=0,\ y=0. \end{cases}$

(b) $f(x,y)=\begin{cases} (x^4+y^4)\log(x^2+y^2) & \text{当 } x,y\neq 0, \\ 0 & \text{当 } x=0,\ y=0. \end{cases}$

d. 微分次序的改变

在第 35–37 页上所给出的全部求偏导数的例子中，我们都有 $f_{yx}=f_{xy}$；换句话说，无论我们是首先对 x 求导，然后对 y 求导，或者是首先对 y 求导，然后对 x 求导，并无差别. 这在下述定理的条件下一般说来是对的.

若一个函数 $f(x,y)$ 的"混合"偏导数 f_{xy} 与 f_{yx} 在开集 R 内连续，则等式

$$f_{yx}=f_{xy} \tag{10}$$

在整个 R 内成立，即与对 x 求导和对 y 求导的次序无关.

证明可像上一小节一样，在微分学中值定理的基础上得到. 我们来考虑四个点 (x,y)，$(x+h,y)$，$(x,y+k)$，和 $(x+h,y+k)$，其中 $h\neq 0$ 和 $k\neq 0$. 如果 (x,y) 是开集 R 内的一个点并且 h 和 k 足够小，则所有这些点均属于 R. 我们现在作表示式

$$A=f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y). \tag{11}$$

引进变量 x 的函数

$$\phi(x)=f(x,y+k)-f(x,y),$$

并把变量 y 视为一个"参数". 于是 A 获得以下形式：

$$A=\phi(x+h)-\phi(x).$$

应用微分学中值定理即得

$$A = h\phi'(x + \theta h),$$

其中 θ 介于 0 到 1 之间，但是根据 $\phi(x)$ 的定义，我们有

$$\phi'(x) = f_x(x, y + k) - f_x(x, y).$$

又因我们已假设“混合”二阶偏导数 f_{yx} 是存在的，所以我们可再次引用中值定理得到

$$A = hkf_{yx}(x + \theta h, y + \theta' k),$$

其中 θ 和 θ' 表示两个非特定的介于 0 到 1 之间的数.

用完全相同的方法，我们可以引进函数

$$\psi(y) = f(x + h, y) - f(x, y)$$

并把 A 表示为

$$A = \psi(y + k) - \psi(y).$$

这样我们就得到等式

$$A = hkf_{xy}(x + \theta_1 h, y + \theta_1' k),$$

其中 $0 < \theta_1 < 1$ 和 $0 < \theta_1' < 1$. 我们令两个 A 的表示式相等，就得到等式

$$f_{yx}(x + \theta h, y + \theta' k) = f_{xy}(x + \theta_1 h, y + \theta_1' k).$$

如果我们在这里令 h 和 k 同时趋向于 0，并记住导数 $f_{xy}(x, y)$ 和 $f_{yx}(x, y)$ 均在点 (x, y) 处连续，我们立即得到

$$f_{yx}(x, y) = f_{xy}(x, y),$$

这正是所要证明的[1].

关于微分次序的可换性定理 (即关于微分算子 D_x 和 D_y 的交换性) 还有更广泛的结论. 特别是我们看到, 多元函数的二阶导数和高阶导数的互不相同的个数比我们原先所想像的要少得多. 如果我们假设我们要计算的导数在我们考虑的区域内都是自变数的连续函数, 并且用函数 $f_x(x,y), f_y(x,y), f_{xy}(x,y)$ 等等去替代函数 $f(x,y)$, 那么引用我们的定理就可得到等式

$$f_{xxy} = f_{xyx} = f_{yxx},$$

$$f_{xyy} = f_{yxy} = f_{yyx},$$

$$f_{xxyy} = f_{xyxy} = f_{xyyx} = f_{yxxy} = f_{yxyx} = f_{yyxx},$$

等等; 一般地说我们有下述结果:

对二元函数进行重复求导时, 求导的次序可以任意改变, 只

1) 作为进一步精细的研究, 有必要知道关于微分次序可换性定理可以在较弱的假设下得到证明. 事实上, 除了假设一阶偏导数 f_x 和 f_y 存在外, 只要还有一个混合偏导数例如 f_{yx} 存在, 并且这个导数在所考虑的点处连续就足够了. 要证明这个结果, 我们回到方程 (11), 先除以 hk, 并令 k 单独趋向于 0, 则右边有极限, 因此左边也有极限, 并且

$$\lim_{k\to 0}\frac{A}{hk} = \frac{f_y(x+h,y)-f_y(x,y)}{h}.$$

更进一步, 我们曾经只用了 f_{yx} 存在这一个假设, 就证明了

$$\frac{A}{hk} = f_{yx}(x+\theta h, y+\theta' k).$$

由于假定了 f_{yx} 连续, 对任意 $\varepsilon > 0$ 与所有足够小的 h 和 k, 我们有

$$f_{yx}(x,y) - \varepsilon < f_{yx}(x+\theta h, y+\theta' k) < f_{yx}(x,y) + \varepsilon,$$

由此得出

$$f_{yx}(x,y) - \varepsilon \le \frac{f_y(x+h,y)-f_y(x,y)}{h} \le f_{yx}(x,y) + \varepsilon,$$

或

$$\lim_{h\to 0}\frac{f_y(x+h,y)-f_y(x,y)}{h} = f_{yx}(x,y),$$

这就是

$$f_{xy}(x,y) = f_{yx}(x,y).$$

要所涉及的导数均为连续函数.[1)]

用我们关于连续性的假设，二元函数有三个二阶偏导数，

$$f_{xx},\quad f_{xy},\quad f_{yy};$$

四个三阶偏导数，

$$f_{xxx},\quad f_{xxy},\quad f_{xyy},\quad f_{yyy};$$

而一般地说，有 $(n+1)$ 个 n 阶偏导数，

$$f_{x^n}, f_{x^{n-1}y}, f_{x^{n-2}y^2}, \cdots, f_{xy^{n-1}}, f_{y^n}.$$

显然，对于两个以上自变量的函数也有类似的陈述. 因为，我们可以同样应用我们的证明到对 x 和 z，或对 y 和 z 等等的求导交换次序上，因为每回相继两次求导时，都只牵涉到两个自变量.

练 习 1.4 d

1. 求出 $\partial^2 z/(\partial x\partial y)$ 和 $\partial^2 z/(\partial y\partial x)$ 并验证它们相等.

(a) $z=(ax+by)^2$, (b) $z=\sqrt{ax+by}$,

(c) $z=f(ax+by)$, (d) $z=ye^x$,

(e) $z=\log\dfrac{x+y}{x}$, (f) $z=e^{\cos(y^2+x)}$.

---

1) 很有兴趣的是用一个例子来表明，不假设二阶导数 f_{xy} 和 f_{yx} 的连续性，定理就不再成立，且 f_{xy} 可以与 f_{yx} 不同. 这可以举下列函数为例：

$$f(x,y)=xy\frac{x^2-y^2}{x^2+y^2},\quad f(0,0)=0.$$

对此所有二阶偏导数均存在但并不连续，我们可得

$$f_x(0,y)=\lim_{x\to 0}\frac{f(x,y)-f(0,y)}{x}=\lim_{x\to 0}y\frac{x^2-y^2}{x^2+y^2}=-y,$$

$$f_y(x,0)=\lim_{y\to 0}\frac{f(x,y)-f(x,0)}{y}=\lim_{y\to 0}x\frac{x^2-y^2}{x^2+y^2}=x,$$

结果

$$f_{yx}(0,0)=-1 \text{ 而 } f_{xy}(0,0)=+1.$$

这两个表示式不同. 根据上述定理，只能是由于 f_{xy} 在原点不连续引起的.

2. 求下列函数的所有三阶偏导数:

(a) $f(x,y)=x^y$,

(b) $f(x,y)=\cosh xy$,

(c) $f(x,y)=ax^2+bxy+cy^2$,

(d) $f(x,y)=\dfrac{x}{y}+\dfrac{y}{x}$,

(e) $f(x,y)=2\cos x+3\sin(y-x)$.

3. 对于函数 $f(x,y)=\log(e^x+e^y)$ 证明

$$f_x+f_y=1 \text{ 以及 } f_{xx}f_{yy}-(f_{xy})^2=0.$$

问 题 1.4 d

1. (a) 证明形如 $u(x,y)=f(x)g(y)$ 的函数满足偏微分方程

$$uu_{xy}-u_xu_y=0;$$

(b) 证明这一断言的逆.

2. 定义 $f(x,y)$ 如下:

$$f(x,y)=\begin{cases} x^2\arctan\dfrac{y}{x}-y^2\arctan\dfrac{x}{y} & \text{当 } x,y\neq 0,\\ 0 & \text{当 } x=0,y=0,\end{cases}$$

求证 $f_{xy}(0,0)=-1$, $f_{yx}(0,0)=1$.

1.5 函数的全微分及其几何意义

a. 可微性的概念

对于一元函数 $y=f(x)$, 导数的存在是与在一个 x 值的邻域内用一个线性函数来近似表示函数 f 的可能性密切相关的; 在几何上, 这相当于把 f 的图形用其切线近似代替. 按照定义, 如果极限

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=A$$

存在，则函数 f 在点 x 处有导数；极限值 A 用 $f'(x)$ 表示. 于是，f 在 x 点处的可微性就意味着，对固定的 x，自变量的增量 $h=\Delta x$ 所对应的增量 $\Delta f=f(x+h)-f(x)$ 可表成这样的形式

$$\Delta f=f(x+h)-f(x)=Ah+\varepsilon h, \tag{12}$$

其中 A 不依赖于 h 而且 $\lim\limits_{h\to 0}\varepsilon=0$. 令 $x+h=\xi$，我们可以说，$f(\xi)$ 可用 ξ 的一个线性函数，即 $\phi(\xi)=f(x)+A(\xi-x)$ 来近似表示，其误差比 $(\xi-x)$ 的一阶更高：

$$f(\xi)-\phi(\xi)=\varepsilon(\xi-x)=o(\xi-x)\ \text{当}\ \xi\to x.$$

当然，以 ξ,η 为变动坐标的线性函数

$$\eta=\phi(\xi)=f(x)+f'(x)(\xi-x)$$

的图形正好是 f 在点 (x,y) 处的切线. 换一种说法，f 在 x 处的可微性意味着，增量 Δf 看作 $h=\Delta x$ 的函数，它能用线性函数

$$df=f'(x)h=f'(x)dx^{1)}$$

来近似代替，误差比 h 的一阶更高.

以上这些看法可以非常自然地推广到两个或者更多个自变量的函数.

我们说函数 $u=f(x,y)$ 在点 (x,y) 是可微的，如果在这个点的邻域内它能用一个线性函数近似表示，也就是说，如果它能表成这样的形式：

$$f(x+h,y+k)=Ah+Bk+C+\varepsilon\sqrt{h^2+k^2}, \tag{13}$$

其中 A,B 和 C 是与变量 h 和 k 无关的，并且 ε 随 h 和 k 趋于 0 而趋于 0. 换句话说，在点 $(x+h,y+k)$ 处函数 $f(x+h,y+k)$ 与

1) 对自变量 x 我们有 $dx=1\cdot h=h=\Delta x$.

h 和 k 的线性函数 $Ah+Bk+C$ 之间的差必须是 $o(\rho)$ 的数量级，这里 $\rho=\sqrt{h^2+k^2}$ 表示点 $(x+h,y+k)$ 到点 (x,y) 的距离.

如果这样一种近似表示是可能的，那么立即可推出，函数 $f(x,y)$ 在点 (x,y) 处是连续的并有关于 x 与 y 的偏导数，而且

$$A=f_x(x,y),\quad B=f_y(x,y),\quad C=f(x,y).$$

首先，从 (13) 可看出，当 $h=k=0$ 时 $f(x,y)=C$. 此外，

$$\lim_{\substack{h\to 0\\ k\to 0}} f(x+h,y+k)=C=f(x,y).$$

所以 f 在点 (x,y) 处是连续的. 在 (13) 中令 $k=0$ 并除以 h，即得关系式

$$\frac{f(x+h,y)-f(x,y)}{h}=A+\varepsilon.$$

由于当 h 趋于 0 时 ε 趋于 0，因而等式左边有极限而且极限就是 A. 类似地，我们得到等式 $f_y(x,y)=B$.

反过来，我们来证明下述基本事实：

一个函数 $u=f(x,y)$ 在刚才定义的意义下是可微的 —— 亦即，它可像 (13) 中那样用一个线性函数近似地表示，误差为 $o(\rho)$ —— 如果它在所考虑的点具有连续的一阶导数.

事实上，我们把函数的增量

$$\Delta u=f(x+h,y+k)-f(x,y)$$

写成这样的形式

$$\begin{aligned}\Delta u=&[f(x+h,y+k)-f(x,y+k)]\\&+[f(x,y+k)-f(x,y)].\end{aligned}$$

正如前面 (第 38 页) 一样，对这两个括号应用微积分学中通常的微分中值定理，即可把 Δu 表成

$$\Delta u=hf_x(x+\theta_1 h,y+k)+kf_y(x,y+\theta_2 k),$$

其中 $0<\theta_1,\theta_2<1$. 因为假设了偏导数 f_x 和 f_y 在点 (x,y) 处连续，我们可以写成

$$f_x(x+\theta_1 h, y+k)=f_x(x,y)+\varepsilon_1$$

和

$$f_y(x, y+\theta_2 k)=f_y(x,y)+\varepsilon_2,$$

其中数 ε_1 与 ε_2 当 h 与 k 趋于 0 时均趋向于 0. 这样我们得到

$$\begin{aligned}\Delta u &= hf_x(x,y)+kf_y(x,y)+\varepsilon_1 h+\varepsilon_2 k\\ &= hf_x(x,y)+kf_y(x,y)+o(\sqrt{h^2+k^2}),\end{aligned}$$

而这个等式表明了 f 的可微性 [1].

有时我们把具有连续的一阶偏导数的函数称为**连续可微的**函数或说它是 C^1 **类**函数. 我们看到, C^1 类的函数都是可微的. 如果除一阶偏导数外所有二阶偏导数也都是连续的, 我们就说这个函数是**二次连续可微的**, 或说它属于 C^2 类, 等等. 连续函数也被称为是 C^0 类函数 [2].

练　习　1.5 a

1. 证明下列每一个函数都在原点不可微:

(a) $f(x,y)=\sqrt{x}\cos y$,

(b) $f(x,y)=\sqrt{|xy|}$,

(c) $f(x,y)=\begin{cases}\dfrac{2xy}{\sqrt{x^2+y^2}}, & 当\ (x,y)\neq(0,0),\\ 0, & 当\ (x,y)=(0,0).\end{cases}$

---

1) 如果我们仅假设导数 f_x 与 f_y 存在但不假设连续, 则函数未必可微 (参见第 37 页).

2) 上面 C^1, C^2 类等等的定义, 仅适用于这样的函数 f, 其定义域是开集, 因为偏导数仅在区域的内点上有定义. 对于定义域 R 是非开集的函数 f 的类, 可以这样推广, 所考虑的 f 的导数在 R 的全部内点上存在并且和一个在整个 R 上都有定义并连续的函数重合.

2. 对于 x, y 分别在区间 $[x_0, x_1], [y_0, y_1]$ 内连续的函数 $g(x)$, $h(y)$, 证明函数

$$f(x,y) = \left(\int_{x_0}^{x} g(s)ds\right) \times \left(\int_{y_0}^{y} h(t)dt\right)$$

当 $x_0 \le x \le x_1$, $y_0 \le y \le y_1$ 时, 在 (x,y) 处是可微的.

问 题 1.5 a

1. 设在点 (a,b) 的一个邻域内,

$$f(x,y) = f(a,b) + hf_x(a,b) + kf_y(a,b) + o(\sqrt{h^2+k^2}),$$

其中 $h = x - a$, $k = y - b$. 假设 f_x, f_y 在点 (a,b) 处存在但在此点未必连续, 求证 f 在 (a,b) 处连续.

b. 方向导数

可微函数 f 的一个基本性质是, 它们不仅具有关于 x 和 y 的偏导数 —— 或者也可以说, 沿 x 轴方向和 y 轴方向的导数 —— 而且还有沿任何方向的导数, 并且这些导数都可以用 f_x 和 f_y 来表示, 所谓**沿 α 方向的导数**是指, 当动点沿着一条与 x 轴的正向夹角为 α 的射线逼近 (x,y) 时, 函数 f 在 (x,y) 处关于距离的变化率. 对于这条射线上的点 $(x+h, y+k)$, h 和 k 可表成

$$h = \rho\cos\alpha, \quad k = \rho\sin\alpha,$$

其中 $\rho = \sqrt{h^2+k^2}$ 是 $(x+h, y+k)$ 到 (x,y) 的距离. 沿着这条射线, f 成了 ρ 的函数

$$f(x+\rho\cos\alpha, y+\rho\sin\alpha).$$

f **在点** (x,y) **处沿 α 方向的导数**定义为 $f(x+\rho\cos\alpha, y+\rho\sin\alpha)$ 在 $\rho = 0$ 处的关于 ρ 的导数, 并且 $D_{(\alpha)}f(x,y)$ 表示. 因此,

$$D_{(\alpha)}f(x,y) = \left(\frac{d}{d\rho}f(x+\rho\cos\alpha, y+\rho\sin\alpha)\right)_{\rho=0}$$

$$= \lim_{\rho \to 0} \frac{f(x + \rho\cos\alpha, y + \rho\sin\alpha) - f(x, y)}{\rho},$$

如果这个极限存在的话. 特别地, 当 $\alpha = 0$ 和 $\alpha = \dfrac{\pi}{2}$ 时, 就得到 f 的两个偏导数:

$$D_{(0)}f(x, y) = \lim_{\rho \to 0} \frac{f(x + \rho, y) - f(x, y)}{\rho} = f_x(x, y),$$

$$D_{\left(\frac{\pi}{2}\right)}f(x, y) = \lim_{\rho \to 0} \frac{f(x, y + \rho) - f(x, y)}{\rho} = f_y(x, y).$$

如果 $f(x, y)$ 是可微的, 则我们有

$$\begin{aligned} f(x + h, y + k) - f(x, y) &= hf_x + kf_y + \varepsilon\rho \\ &= \rho(f_x\cos\alpha + f_y\sin\alpha + \varepsilon). \end{aligned} \tag{14}$$

令 ρ 趋于 0, 于是, 由于 ε 趋于 0, 我们就得到 f 沿 α 方向的导数的表达式

$$D_{(\alpha)}f(x, y) = f_x\cos\alpha + f_y\sin\alpha. \tag{14a}$$

这样, **方向导数** $D_{(\alpha)}f$ **便是** x **轴方向和** y **轴方向的导数** f_x **和** f_y **的线性组合**, 系数为 $\cos\alpha$ **和** $\sin\alpha$. 特别地, 当导数 f_x 和 f_y 在所考虑的点存在而且连续时, 上述结果成立.

例如, 取 $f(x, y)$ 为由原点到点 (x, y) 的距离

$$r = \sqrt{x^2 + y^2},$$

我们有偏导数

$$r_x = \frac{x}{\sqrt{x^2 + y^2}} = \frac{x}{r} = \cos\theta$$

和

$$r_y = \frac{y}{\sqrt{x^2 + y^2}} = \frac{y}{r} = \sin\theta,$$

其中 θ 是矢径与 x 轴构成的夹角. 因而, 在 α 方向上函数 r 具有导数

$$D_{(\alpha)}r = r_x \cos\alpha + r_y \sin\alpha = \cos\theta\cos\alpha + \sin\theta\sin\alpha$$
$$= \cos(\theta - \alpha);$$

特别地, 沿矢径本身的方向 (即从原点直达 (x,y) 的方向), 这个导数具有数值 1, 而沿与矢径垂直的方向, 它具有数值 0.

对函数 x, 沿着矢径的方向, 有导数 $D_\theta(x) = \cos\theta$, 而对函数 y, 有导数 $D_\theta(y) = \sin\theta$; 沿着与矢径垂直的方向, 这两个函数的导数则分别为

$$D_{(\theta+\pi/2)}x = -\sin\theta \text{ 和 } D_{(\theta+\pi/2)}y = \cos\theta.$$

函数 $f(x,y)$ 沿矢径方向的导数一般用 $\partial f(x,y)/\partial r$ 表示. 它实际上是将函数 $f(r\cos\theta, r\sin\theta)$ 看作 r 和 θ 的函数时, 关于 r 的偏导数. 这样, 我们有关系式

$$\frac{\partial f}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y},$$

我们可把这个关系式简便地写成符号形式

$$\frac{\partial}{\partial r} = \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y},$$

这如同微分算子 $\partial/\partial r$, $\partial/\partial x$, $\partial/\partial y$ 之间的恒等式.

值得注意的是, 如果坐标为 $(x+h, y+k)$ 的点 Q 不是沿着方向为 α 的直线逼近坐标为 (x,y) 的点 P 而是沿任一条曲线逼近 P, 而该曲线在 P 点的切线方向为 α, 这样我们也可以得到函数 $f(x,y)$ 在 α 方向上的导数. 因为如果直线 PQ 有方向 β, 我们可写 $h = \rho\cos\beta$, $k = \rho\sin\beta$, 并在上面证明过程中所用的公式 (14) 中把 α 换成 β. 但由于根据假设当 $\rho \to 0$ 时 β 趋向于 α, 我们最后仍得到关于 $D_{(\alpha)}f(x,y)$ 的同样的关系式.

用同样的方法，一个有三个自变量的可微函数 $f(x,y,z)$ 也可在给定的方向求导数. 我们假设这个方向由它与坐标轴所构成的三个角度的余弦所规定. 若我们把这三个角度叫做 α,β,γ 并且考虑两个点 (x,y,z) 和 $(x+h,y+k,z+l)$, 其中

$$h=\rho\cos\alpha,\quad k=\rho\cos\beta,\quad l=\rho\cos\gamma,$$

于是像 (14a) 一样，我们得到在角度 (α,β,γ) 所确定的方向上的导数表达式

$$f_x\cos\alpha+f_y\cos\beta+f_z\cos\gamma. \tag{14b}$$

练 习 1.5 b

1. 函数 f 在倾角 α 确定的方向上的导数 $D_{(\alpha)}f(x,y)$ 的几何解释是什么？

2. 对下列函数，求出 $D_{(\alpha)}f(x,y)$, $\alpha=0°, 30°, 60°, 90°$:

(a) $f(x,y)=ax+by$, a,b 为常数, $x_0=y_0=0$.

(b) $f(x,y)=ax^2+by^2$, $x_0=y_0=1$ (a,b 为常数).

(c) $f(x,y)=x^2-y^2$, $x_0=1$, $y_0=2$.

(d) $f(x,y)=\sin x+\cos y$, $x_0=y_0=0$.

(e) $f(x,y)=e^x\cos y$, $x_0=0$, $y_0=\pi$.

(f) $f(x,y)=\sqrt{2x^2+y^2}$, $x_0=1$, $y_0=1$.

(g) $f(x,y)=\cos(x+y)$, $x_0=0$, $y_0=0$.

3. 按照所给规定求出下列每一个函数的方向导数:

(a) $z^2-x^2-y^2$ 在点 $(1,0,1)$, 沿方向 $(4,3,0)$.

(b) $xyz-xy-yz-zx+x+y+z$ 在点 $(2,2,1)$ 沿方向 $(2,2,0)$.

(c) $xz^2+y^2+z^3$ 在点 $(1,0,-1)$ 沿方向 $(2,1,0)$.

4. 给出这样一个函数的例子，它在一个点处沿每一个方向都有导数，可是在该点是不可微的.

5. 对于 $f(x,y)=\sqrt[3]{xy}$, 证明 f 在原点是连续的而且偏导数 $\partial z/\partial x$ 和 $\partial z/\partial y$ 存在, 但沿其他所有方向的方向导数都不存在.

6. 设 $f(x,y)=xy+\sqrt{2x^2+y^2}$, $r=\sqrt{x^2+y^2}$, $y/x=\operatorname{tg}\theta$. 对 $\theta=0°, 30°, 60°, 90°$, 与 $x,y=1$, 求出 $\partial^2 f/\partial r^2$.

c. 可微性的几何解释. 切平面

对于一个函数 $z=f(x,y)$, 所有这些概念都很容易从几何上来说明. 我们回顾关于 x 的偏导数是这样一条曲线的切线的斜率, 这条曲线是表示关系式 $z=f(x,y)$ 的曲面与一个垂直于 $x-y$ 平面且平行于 x 轴的平面相交而成的. 同样的道理, 沿 α 方向的导数给出了这个曲面与一个通过 (x,y,z) 、垂直于 xy 平面且与 x 轴的夹角是 α 的平面相交而成的曲线的切线的斜率. 这样一来, 公式

$$D_{(\alpha)}f(x,y)=f_x\cos\alpha+f_y\sin\alpha$$

使我们能算出所有这样的曲线的切线的斜率, 也就是, 从两条这样的切线的斜率, 能算出曲面上一给定点处的所有切线的斜率 [1].

我们曾经把可微函数 $\zeta=f(\xi,\eta)$ 在点 (x,y) 的邻域内用线性函数

$$\phi(\xi,\eta)=f(x,y)+(\xi-x)f_x+(\eta-y)f_y$$

逼近, 其中 ξ 和 η 是变动坐标. 在几何上, 这个线性函数由一个平面表示, 与曲线的切线类似, 我们称这个平面为曲面的 **切平面**. 当

---

1) 对该平面上的每一点 (ξ,η,ζ), 我们有 $\xi=x+\rho\cos\alpha$, $\eta=y+\rho\sin\alpha$, 因而对相截而得的曲线上的点, 我们有

$$\zeta=f(x+\rho\cos\alpha, y+\rho\sin\alpha).$$

以 ρ 与 ζ 为坐标, 这个曲线在 $\zeta=z$, $\rho=0$ 处的切线的斜率由下式给出

$$\left(\frac{d\zeta}{d\rho}\right)_{\rho=0}=D_{(\alpha)}f(x,y).$$

因而, 切线方程是

$$\zeta=z+\rho D_{(\alpha)}f(x,y)=f(x,y)+\rho\cos\alpha f_x(x,y)+\rho\sin\alpha f_y(x,y).$$

$\xi - x = h$ 和 $\eta - y = k$ 趋于 0 时，这个线性函数与函数 $f(\xi,\eta)$ 之间的差比 $\sqrt{h^2+k^2}$ 更高阶地趋于零. 回顾平面曲线切线的定义，这意味着，这个切平面与任何一个与 xy 平面垂直的平面相交所得的直线正好是对应的相交曲线的切线. **于是我们看到，曲面在点 (x,y,z) 处的所有切线全都落在同一个平面，即切平面上.**

这个性质是函数在点 (x,y,z) 处的可微性的几何表示，其中 $z=f(x,y)$. 在变动坐标系 (ξ,η,ζ) 中，点 (x,y,z) 处的切平面方程是

$$\zeta - z = (\xi - x)f_x + (\eta - y)f_y.$$

正如在第 46 页中已指出的，只要偏导数在给定点是连续的，函数就在给定点处是可微的. 与一元函数的情况不同，仅有偏导数 f_x 和 f_y 的**存在性**不足以保证函数的可微性. 如果在所考虑的点处导数不连续，那么曲面在该点处的切平面可以不存在；或者，从分析上说，$f(x+h,y+k)$ 与 h 和 k 的线性函数 $f(x,y)+hf_x(x,y)+kf_y(x,y)$ 之间的差不是比 $\sqrt{h^2+k^2}$ 更高阶地趋于 0. 这可以用一个简单的例子清楚地说明：

$$u = f(x,y) = \frac{xy}{\sqrt{x^2+y^2}}, \quad 当\ x^2+y^2 \neq 0,$$
$$u = 0, \qquad 当\ x=0,\ y=0.$$

如果我们引用极坐标，这就变成

$$u = \frac{r}{2}\sin 2\theta.$$

u 关于 x 和 y 的一阶导数在原点的邻域内处处存在而且在原点本身取 0 值. 然而，这些导数在原点是不连续的，因为

$$u_x = y\left(\frac{1}{\sqrt{x^2+y^2}} - \frac{x^2}{\sqrt{(x^2+y^2)^3}}\right) = \frac{y^3}{\sqrt{(x^2+y^2)^3}}.$$

如果我们沿着 x 轴逼近原点，u_x 趋于 0；而如果我们沿着 y 轴逼近，u_x 就趋向于 1. 这个函数在原点是不可微的；在该点处曲面

$z=f(x,y)$ 的切平面不存在. 因为等式 $f_x(0,0)=f_y(0,0)=0$ 表明切平面应当与平面 $z=0$ 重合. 但对于直线 $\theta=\dfrac{\pi}{4}$ 上的所有点, 我们有

$$\sin 2\theta=1 \text{ 和 } z=f(x,y)=\frac{r}{2};$$

这样, 从这个平面上的点到曲面上的点的距离 z, 并不比 r 更高阶地趋于 0, 而这却是切平面上的点必须具备的条件. 这个曲面是一个以原点为顶点的锥面, 它的母线并不全落在同一个平面上.

练 习 1.5 c

1. 对下列每一个情况, 求出由 $z=f(x,y)$ 定义的曲面在点 $P=(x_0,y_0)$ 处的切平面方程:

(a) $f(x,y)=3x^2+4y^2,\quad P=(0,1)$.

(b) $f(x,y)=2\cos(x-y)+3\sin x,\quad P=\left(\pi,\dfrac{\pi}{2}\right)$.

(c) $f(x,y)=\cosh(x+y),\quad P=(0,\log 2)$.

(d) $f(x,y)=\sqrt{x^2+y^2},\quad P=(1,2)$.

(e) $f(x,y)=e^{x\cos y},\quad P=\left(1,\dfrac{\pi}{4}\right)$.

(f) $f(x,y)=\cos\pi e^{xy},\quad P=(\log 2,1)$.

(g) $f(x,y)=\displaystyle\int_0^{x^2+y^2}e^{-t^2}dt,\quad P=(1,1)$.

(h) $f(x,y)=ax^3+bx^2y+cxy^2+dy^3,\quad P=(1,1)$,

(a,b,c,d 为常数).

2. 证明曲面 $z=yf(x/y)$ 的全部切平面相交于一个公共点, 其中 f 是任何一个可微的一元函数.

3. 说明曲面 $S:z=f(x,y)$ 在点 $P_0(x_0,y_0)$ 处的切平面是通过 S 上三个点 (x_i,y_i,z_i), $i=0,1,2$ 的平面的极限位置, 其中 $P_1=(x_1,y_1)$ 和 $P_2=(x_2,y_2)$ 沿着角度不等于 0° 或 180° 的不同的方向逼近于 P_0.

4. 证明二次曲面

$$ax^2+by^2+cz^2=1$$

在点 (x_0, y_0, z_0) 处的切平面是

$$ax_0x + by_0y + cz_0z = 1.$$

d. 函数的微分

像对一元函数一样，对出现在公式 (14)

$$\begin{aligned}\Delta u &= f(x+h, y+k) - f(x, y)\\ &= hf_x(x, y) + kf_y(x, y) + \varepsilon\sqrt{h^2+k^2}\end{aligned}$$

中的可微函数 $u = f(x, y)$ 的增量的线性部分，有一个专门的名字与符号往往是方便的．我们称这个线性部分为函数的**微分**，并写成

$$\begin{aligned}du = df(x, y) &= \frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k\\ &= \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y.\end{aligned} \tag{15a}$$

这个微分，有时称为**全微分**，是**四**个自变量，即，所考虑的点的坐标 x 与 y 以及自变量的增量 h 与 k 的函数．我们再次强调，这里没有含糊的"无穷小量"的概念．这仅仅意味着： du 是函数的增量

$$\Delta u = f(x+h, y+k) - f(x, y)$$

的近似表示式，其误差是 $\sqrt{h^2+k^2}$ 的 ε 倍， ε 可任意小，只要 h 和 k 是充分小的量．对于自变量 x 和 y，从 (15a) 我们得出

$$dx = \frac{\partial x}{\partial x}\Delta x + \frac{\partial x}{\partial y}\Delta y = \Delta x,$$

$$dy = \frac{\partial y}{\partial x}\Delta x + \frac{\partial y}{\partial y}\Delta y = \Delta y.$$

因此，微分 $df(x, y)$ 更通常写成

$$df(x, y) = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = f_x(x, y)dx + f_y(x, y)dy. \tag{15b}$$

顺便说一下，微分完全确定了 f 的一级偏导数. 例如. 当取 $dy=0$ 与 $dx=1$ 时，我们从 df 就能得到偏导数 $\partial f/\partial x$.

我们强调指出，除非函数在上面定义的意义下是可微的 (对这一点，两个偏导数的连续性，而不仅仅是存在，就足够了)，函数 $f(x,y)$ 的全微分作为 Δf 的线性近似式是没有意义的.

如果函数 $f(x,y)$ 还有更高阶的连续偏导数，我们就能建立微分 $df(x,y)$ 的微分；亦即，我们将它的关于 x 和 y 的偏导数分别乘以 $h=dx$ 和 $k=dy$, 然后将这些乘积加起来. 在这个微分法中，我们将 h 和 k 当作常数，这相应于微分

$$df = hf_x(x,y) + kf_y(x,y)$$

是四个自变量 x,y,h 和 k 的函数这一事实. 于是我们得到函数的**二阶微分**[1]

$$\begin{aligned} d^2 f = d(df) &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k\right)h + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k\right)k \\ &= \frac{\partial^2 f}{\partial x^2}h^2 + 2\frac{\partial^2 f}{\partial x\partial y}hk + \frac{\partial^2 f}{\partial y^2}k^2 \\ &= \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x\partial y}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2. \end{aligned}$$[2]

类似地，我们可建立**更高阶的微分**

$$\begin{aligned} d^3 f = d(d^2 f) &= \frac{\partial^3 f}{\partial x^3}dx^3 + 3\frac{\partial^3 f}{\partial x^2\partial y}dx^2dy \\ &\quad + 3\frac{\partial^3 f}{\partial x\partial y^2}dxdy^2 + \frac{\partial^3 f}{\partial y^3}dy^3, \\ d^4 f &= \frac{\partial^4 f}{\partial x^4}dx^4 + 4\frac{\partial^4 f}{\partial x^3\partial y}dx^3dy + 6\frac{\partial^4 f}{\partial x^2\partial y^2}dx^2dy^2 \end{aligned}$$

---

1) 我们将在以后 (第 76 页) 看到，这里从形式上引进的高阶微分，正好与函数展开式中同阶的项对应.

2) 习惯上，微分的幂 $(dx)^2,(dx)^3,(dy)^2,(dy)^3$ 简写成 dx^2,dx^3,dy^2,dy^3. 这样，当然，有些会引起误会，因为它们可能会与 $d(x^2)=2xdx$, $d(x^3)=3x^2dx$ 等等发生混淆.

$$+4\frac{\partial^4 f}{\partial x\partial y^3}dxdy^3+\frac{\partial^4 f}{\partial y^4}dy^4,$$

并且，用归纳法容易证明，一般地有

$$d^n f=\frac{\partial^n f}{\partial x^n}dx^n+\binom{n}{1}\frac{\partial^n f}{\partial x^{n-1}\partial y}dx^{n-1}dy+\cdots$$
$$+\binom{n}{k}\frac{\partial^n f}{\partial x^{n-k}\partial y^k}dx^{n-k}dy^k+\cdots+\frac{\partial^n f}{\partial y^n}dy^n.$$

最后这个公式可用符号表示为方程

$$d^n f=\left(\frac{\partial}{\partial x}dx+\frac{\partial}{\partial y}dy\right)^n f,$$

这里右端的表达式可理解为先根据二项式定理形式地展开，然后以

$$\frac{\partial^n f}{\partial x^n}dx^n,\ \frac{\partial^n f}{\partial x^{n-1}\partial y}dx^{n-1}dy,\cdots,\ \frac{\partial^n f}{\partial y^n}dy^n$$

去代替

$$\left(\frac{\partial}{\partial x}dx\right)^n f,\ \left(\frac{\partial}{\partial x}dx\right)^{n-1}\left(\frac{\partial}{\partial y}dy\right)f,\cdots,\left(\frac{\partial}{\partial y}dy\right)^n f.$$

对微分的计算，法则

$$d(fg)=fdg+gdf$$

仍然成立；这可由乘积的微分法立即得出.

最后，我们注意，这一节的讨论可直接推广到多于两个自变量的函数.

练 习 1.5 d

1. 求出下列函数的全微分：

(a) $z=x^2y^2+3xy^3-2y^4$,

(b) $z=\dfrac{xy}{x^2+2y^2}$,

(c) $z=\log(x^4-y^3)$,

(d) $z=\dfrac{x}{y}+\dfrac{y}{x}$,

(e) $z=\cos(x+\log y)$,

(f) $z=\dfrac{x-y}{x+y}$,

(g) $z=\arctan(x+y)$,

(h) $z=x^y$,

(i) $w=\cosh(x+y-z)$,

(j) $w=x^2-2xz+y^3$.

2. 对 $x=1$, $y=2$, $dx=0.1$, $dy=0.3$ 计算 $f(x,y)=x-y+(x^2+y^2)^{1/3}$ 的全微分的值.

3. 对 $f(x,y)=e^{x^2+y^2}$ 求 $d^3f(x,y)$.

e. 在误差计算方面的应用

在实用中，微分 $df=hf_x+kf_y$ 经常被用来作为函数 $f(x,y)$ 从 (x,y) 到 $(x+h,y+k)$ 的增量 $\Delta f=f(x+h,y+k)-f(x,y)$ 的近似表达式. 这个用途在所谓 “误差计算” 中特别充分地显示出来了 (参看第一卷第 554 页). 例如，假设我们要求用排量法测求固体的密度时所可能引起的误差. 若 m 是物体在空气中的重量，$\overline{m}$ 是它浸没在水中时的重量，则根据阿基米德原理，重量的亏损 $(m-\overline{m})$ 是被排出的水的重量. 如果我们用 cgs (厘米 - 克 - 秒) 单位制，被排出的水的重量在数值上就等于它的体积，因而也等于固体的体积. 所以这个固体的密度 s 是由公式 $s=m/(m-\overline{m})$ 给出，其中 m 和 $\overline{m}$ 作为自变量. 由 m 的测量误差 dm 和 $\overline{m}$ 的测量误差 $d\overline{m}$ 所引起的密度 s 的测量误差可由全微分

$$ds=\frac{\partial s}{\partial m}dm+\frac{\partial s}{\partial \overline{m}}d\overline{m}$$

近似表示. 根据 (求导数的) 除法法则，偏导数是

$$\frac{\partial s}{\partial m}=-\frac{\overline{m}}{(m-\overline{m})^2} \text{ 和 } \frac{\partial s}{\partial \overline{m}}=\frac{m}{(m-\overline{m})^2};$$

所以这个微分是

$$ds = \frac{-\overline{m}dm + md\overline{m}}{(m-\overline{m})^2}.$$

这样，如果 dm 和 $d\overline{m}$ 有相反的符号，s 的误差就最大，譬如说，如果我们以一个过小的数 $m+dm$ 来代替 m，又以一个过大的数 $\overline{m}+d\overline{m}$ 来代替 $\overline{m}$. 例如，若有一黄铜片，在空气中的重量约为 100g，可能的误差为 0.005g，而在水中的重量约为 88g，可能的误差为 0.008g，则由我们的公式所给出的密度的误差大约是

$$\frac{88\cdot 5\cdot 10^{-3}+100\cdot 8\cdot 10^{-3}}{12^2}\sim 9\cdot 10^{-3},$$

即大约是 1%.

练　习　1.5 e

1. 求出函数 $z=(x+y)/(x-y)$ 的近似变差，当 x 从 $x=2$ 变到 $x=2.5$，并且 y 从 $y=4$ 变到 $y=4.5$ 时.

2. 求 $\log[(1.02)^{1/4}+(0.96)^{1/6}-1]$ 的近似值.

3. 已知直角三角形的底长 x 和高 y 分别包含误差 h,k. 面积的可能的误差是什么？

4. 若 dz 是量 z 的测量误差，相对误差定义为 dz/z. 证明，乘积 $z=xy$ 的相对误差是因子的相对误差之和.

5. 重力加速度 g 是以一个下落物体的降落时间 (以秒计算) 来测定的，这个物体从静止开始降落并经过一个固定的距离 x. 若所测时间是 t，我们有 $g=2x/t^2$. 如果 x 约为 1 米和 t 约为 0.45 秒. 说明 g 的相对误差，对于 t 的相对误差比起对于 x 的相对误差来说，更敏感些.

1.6　函数的函数 (复合函数) 与新自变量的引入

a. 复合函数. 链式法则

自变量 x,y 的一个函数 u 常常由形式

$$u = f(\xi, \eta, \cdots)$$

表出，其中 f 的自变量 $\xi, \eta, \cdots$ 本身又是 x 和 y 的函数

$$\xi = \varphi(x, y), \quad \eta = \psi(x, y), \cdots.$$

于是，我们称

$$u = f(\xi, \eta, \cdots) = f(\varphi(x, y), \psi(x, y), \cdots) = F(x, y) \tag{16}$$

是 x 和 y 的一个**复合函数**(参看第一卷第 56 页). 例如，函数

$$u = F(x, y) = e^{xy} \sin(x + y) \tag{16a}$$

可以写成一个复合函数. 为此只需利用关系式

$$u = f(\xi, \eta) = e^{\xi} \sin \eta, \tag{16b}$$

其中 $\xi = xy$ 和 $\eta = x + y$. 类似地，函数

$$u = F(x, y) = \log(x^4 + y^4) \cdot \arcsin \sqrt{1 - x^2 - y^2} \tag{16c}$$

能表成复合函数的形式：

$$u = f(\xi, \eta) = \eta \cdot \arcsin \xi, \tag{16d}$$

其中 $\xi = \sqrt{1 - x^2 - y^2}$ 和 $\eta = \log(x^4 + y^4)$.

为了使复合函数的概念有意义，我们假设函数 $\xi = \varphi(x, y), \eta = \psi(x, y), \cdots$ 有公共的定义域 R, 并且把 R 中的任一点 (x, y) 映射成 $(\xi, \eta, \cdots)$, 而在 $(\xi, \eta, \cdots)$ 上函数 $u = f(\xi, \eta, \cdots)$ 有定义，也就是说，R 中任一点 (x, y) 映射成 f 的定义域 S 中的点. 于是，复合函数

$$u = f(\varphi(x, y), \psi(x, y), \cdots) = F(x, y)$$

定义在区域 R 中.

对区域 R 和 S 的详细研究通常是不必要的，像在 (16b) 中，其中自变量的点 (ξ, η) 能通过整个 $\xi\eta$ 平面因而函数 $u = e^{\xi} \sin \eta$ 定义在

整个 $\xi\eta$ 平面上. 另一方面, (16d) 表明在复合函数的定义中研究区域 R 和 S 的必要性. 因为函数 $\xi=\sqrt{1-x^2-y^2}$ 和 $\eta=\log(x^4+y^4)$ 只在由点 $0<x^2+y^2\leq 1$ 组成的区域 R 内有定义, 亦即在以原点为圆心并将原点除外的闭单位圆内有定义. 在此区域内我们有 $|\xi|<1$, $\eta\leq 0$. 对应的点全都落在函数 $\eta\arcsin\xi$ 的定义域内, 因此复合函数定义在 R 中.

连续函数的一个连续函数仍是连续的. 更确切地说, **如果函数** $u=f(\xi,\eta,\cdots)$ **在区域** S **内是连续的, 并且函数** $\xi=\varphi(x,y)$, $\eta=\psi(x,y),\cdots$ **在区域** R **内是连续的, 那么复合函数** $u=F(x,y)$ **在** R **内是连续的.**

证明可从连续性的定义直接得到. 设 (x_0,y_0) 是 R 中的一个点, 并设 $\xi_0,\eta_0,\cdots$ 是对应的 $\xi,\eta,\cdots$ 的值. 现在, 对任何一个正数 ϵ, 差

$$f(\xi,\eta,\cdots)-f(\xi_0,\eta_0,\cdots)$$

的绝对值都必定小于 ϵ, 只要满足不等式

$$\sqrt{(\xi-\xi_0)^2+(\eta-\eta_0)^2+\cdots}<\delta,$$

这里 δ 是一个足够小的正数. 但由 $\varphi(x,y),\psi(x,y),\cdots$ 的连续性, 上述不等式是必定能满足的, 只要

$$\sqrt{(x-x_0)^2+(y-y_0)^2}<r,$$

这里 r 是一个足够小的正的量. 这就建立了复合函数的连续性.

类似地, **可微函数的一个可微函数仍是可微的**. 这个陈述在下述定理中将更严格地叙述, 定理同时给出了复合函数微分法的法则, 即所谓**链法则**:

如果 $\xi=\varphi(x,y),\eta=\psi(x,y),\cdots$ **是** (x,y) **在区域** R **内的可微函数, 又如果** $f(\xi,\eta,\cdots)$ **是** $(\xi,\eta,\cdots)$ **在区域** S **内的可微函数, 则**

复合函数

$$u = f(\varphi(x,y), \psi(x,y), \cdots) = F(x,y) \tag{17}$$

也是 (x,y) 的可微函数;它的偏导数由公式

$$F_x = f_\xi \varphi_x + f_\eta \psi_x + \cdots, \tag{18}$$
$$F_y = f_\xi \varphi_y + f_\eta \psi_y + \cdots,$$

给出, 或者简写成

$$u_x = u_\xi \xi_x + u_\eta \eta_x + \cdots,$$
$$u_y = u_\xi \xi_y + u_\eta \eta_y + \cdots, \tag{19}$$

这样, 为了求出关于 x 的偏导数, 我们必须首先求出复合函数关于每一个变量 $\xi, \eta, \cdots$ 的导数, 再把这些导数中的每一个乘以对应的变量关于 x 的导数, 并将所有的乘积相加. 这是在第一卷 (第 244 页) 中所讨论过的一元函数的锁链法则的推广.

如果用微分记号, 我们的叙述可写成一种特别简单而又有启发性的形式:

$$\begin{aligned} du &= u_\xi d\xi + u_\eta d\eta + \cdots \\ &= u_\xi(\xi_x dx + \xi_y dy) + u_\eta(\eta_x dx + \eta_y dy) + \cdots \\ &= (u_\xi \xi_x + u_\eta \eta_x + \cdots)dx + (u_\xi \xi_y + u_\eta \eta_y + \cdots)dy \\ &= u_x dx + u_y dy. \end{aligned} \tag{20}$$

这个等式表明, 我们得到复合函数 $u = f(\xi, \eta, \cdots) = F(x,y)$ 的增量的线性部分的步骤是, 先将 $\xi, \eta, \cdots$ 看作自变量而写出这个线性部分, 然后用函数 $\xi = \varphi(x,y)$, $\eta = \psi(x,y), \cdots$ 的增量的线性部分去替换 $d\xi, d\eta, \cdots$. 这个事实显示了微分记号的方便与灵活性.

为了证明法则 (18), 只需利用各有关的函数是可微的假设. 根据这个假设得到, 对应于自变量 x 和 y 的增量 Δx 和 Δy, 量 $\xi, \eta, \cdots$ 的改变量是

$$\Delta\xi = \xi_x \Delta x + \xi_y \Delta y + \varepsilon_1 \sqrt{(\Delta x)^2 + (\Delta y)^2}, \tag{20a}$$

$$\Delta\eta = \eta_x\Delta_x + \eta_y\Delta y + \varepsilon_2\sqrt{(\Delta x)^2 + (\Delta y)^2}, \cdots, \qquad (20b)$$

其中数值 $\varepsilon_1, \varepsilon_2, \cdots$ 当 $\Delta x \to 0$ 和 $\Delta y \to 0$ 或 $\sqrt{(\Delta x)^2 + (\Delta y)^2} \to 0$ 时趋于 0. 导数 $\xi_x, \xi_y, \eta_x, \eta_y$ 是对自变量 x, y 而取的. 此外, 如果量 $\xi, \eta, \cdots$ 改变了 $\Delta\xi, \Delta\eta, \cdots$, 函数 $u = f(\xi, \eta, \cdots)$ 就得到了改变量

$$\Delta u = f_\xi\Delta\xi + f_\eta\Delta\eta + \cdots + \delta\sqrt{(\Delta\xi)^2 + (\Delta\eta)^2 + \cdots}, \qquad (21)$$

其中量 δ 当 $\Delta\xi \to 0$ 和 $\Delta\eta \to 0$ 时趋于 0, f_ξ, f_η 对自变量 ξ, η 而取. 这里, 对 $\Delta\xi, \Delta\eta, \cdots$, 用公式 (20a, b) 所给出的对应于 x, y 的增量 $\Delta x, \Delta y$ 的数值, 我们就得到这样一个等式

$$\begin{aligned}\Delta u =&(f_\xi\xi_x + f_\eta\eta_x + \cdots)\Delta x + (f_\xi\xi_y + f_\eta\eta_y + \cdots)\Delta y\\ &+ \varepsilon\sqrt{(\Delta x)^2 + (\Delta y)^2}.\end{aligned} \qquad (22)$$

这里对于 $\Delta x = \rho\cos\alpha$, $\Delta y = \rho\sin\alpha$, $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$, 量 ε 由下式给出:

$$\begin{aligned}\varepsilon =&\varepsilon_1 f_\xi + \varepsilon_2 f_\eta\\ &+\delta\sqrt{(\xi_x\cos\alpha+\xi_y\sin\alpha+\varepsilon_1)^2+(\eta_x\cos\alpha+\eta_y\sin\alpha+\varepsilon_2)^2+\cdots}.\end{aligned}$$

当 $\rho \to 0$ 时量 $\Delta x, \Delta y, \varepsilon_1, \varepsilon_2$ 趋于 0, 从而 $\Delta\xi, \Delta\eta$ 和 δ 也趋于 0. 另一方面, $f_\xi, f_\eta, \cdots, \xi_x, \xi_y, \eta_x, \eta_y, \cdots$ 是固定不变的, 从而,

$$\lim_{\rho\to 0}\varepsilon = 0.$$

由 (22) 推知, u 作为自变量 x, y 的函数, 在点 (x, y) 处是可微的, 而且 du 由等式 (20) 给出. 由 du 的这个表达式, 我们就知道偏导数 u_x, u_y 具有表达式 (19) 或 (18).

显然这个结果与自变量 $x, y, \cdots$ 的个数无关. 例如, 若量 $\xi, \eta, \cdots$ 仅依赖于一个自变量 x, 从而 u 就是单个变量 x 的一个复合函数, 上述定理仍然是成立的.

为了计算更高阶的偏导数，我们只须将等式 (19) 的右边对 x 和 y 求导数，而把 f_ξ, f_η 作为复合函数处理. 为简单起见，我们仅限于考虑三个函数 ξ,η 和 ζ 的情形，我们得到 [1)]

$$\begin{aligned} u_{xx} =& f_{\xi\xi}\xi_x^2 + f_{\eta\eta}\eta_x^2 + f_{\zeta\zeta}\zeta_x^2 + 2f_{\xi\eta}\xi_x\eta_x \\ &+ 2f_{\eta\zeta}\eta_x\zeta_x + 2f_{\xi\zeta}\xi_x\zeta_x + f_\xi\xi_{xx} + f_\eta\eta_{xx} + f_\zeta\zeta_{xx}, \end{aligned} \tag{23a}$$

$$\begin{aligned} u_{xy} =& f_{\xi\xi}\xi_x\xi_y + f_{\eta\eta}\eta_x\eta_y + f_{\zeta\zeta}\zeta_x\zeta_y + f_{\xi\eta}(\xi_x\eta_y + \xi_y\eta_x) \\ &+ f_{\eta\zeta}(\eta_x\zeta_y + \eta_y\zeta_x) + f_{\xi\zeta}(\xi_x\zeta_y + \xi_y\zeta_x) \\ &+ f_\xi\xi_{xy} + f_\eta\eta_{xy} + f_\zeta\zeta_{xy}, \end{aligned} \tag{23b}$$

$$\begin{aligned} u_{yy} =& f_{\xi\xi}\xi_y^2 + f_{\eta\eta}\eta_y^2 + f_{\zeta\zeta}\zeta_y^2 + 2f_{\xi\eta}\xi_y\eta_y \\ &+ 2f_{\eta\zeta}\eta_y\zeta_y + 2f_{\xi\zeta}\xi_y\zeta_y + f_\xi\xi_{yy} + f_\eta\eta_{yy} + f_\zeta\zeta_{yy}. \end{aligned} \tag{23c}$$

练　习　1.6 a

1. 对下列函数，求出关于 x 和 y 的所有一阶和二阶偏导数:

(a) $z = u\log v$, 其中 $u = x^2$, $v = \dfrac{1}{1+y}$;

(b) $z = e^{uv}$, 其中 $u = ax$, $v = \cos y$;

(c) $z = u\arctan v$, 其中 $u = \dfrac{xy}{x-y}$, $v = x^2y + y - x$;

(d) $z = g(x^2 + y^2, e^{x-y})$;

(e) $z = \tan(x\arctan y)$.

2. 对下列函数计算一阶偏导数:

(a) $w = \dfrac{1}{\sqrt{(x^2 + y^2 + 2xy\cos z)}}$;

(b) $w = \arcsin\dfrac{x}{z + y^2}$;

(c) $w = x^2 + y\log(1 + x^2 + y^2 + z^2)$;

(d) $w = \arctan\sqrt{(x + yz)}$.

3. 计算下列函数的导数:

---

1) 这里假设 f 是 ξ,η,ζ 的 C^2 类函数，而 ξ,η,ζ 是 x,y 的 C^2 类函数. 这就保证了 x 和 y 的复合函数 u 也属于 C^2 类.

(a) $z = x^{(x^x)}$,

(b) $z = \left(\left(\frac{1}{x}\right)^{1/x}\right)^{1/x}$

4. 证明，如果 $f(x,y)$ 满足拉普拉斯方程

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0,$$

那么 $\varphi(x,y) = f\left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2}\right)$ 也如此.

5. 证明函数

(a) $f(x,y) = \log\sqrt{x^2+y^2}$,

(b) $g(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2}}$,

(c) $h(x,y,z,w) = \frac{1}{x^2+y^2+z^2+w^2}$

满足各自的拉普拉斯方程:

(a) $f_{xx} + f_{yy} = 0$,

(b) $g_{xx} + g_{yy} + g_{zz} = 0$,

(c) $h_{xx} + h_{yy} + h_{zz} + h_{ww} = 0$.

问　题　1.6 a

1. 证明，如果 $f(x,y)$ 满足拉普拉斯方程

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0,$$

又如果 $u(x,y)$ 和 $v(x,y)$ 满足柯西 - 黎曼方程

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

那么函数 $\varphi(x,y) = f(u(x,y), v(x,y))$ 也是拉普拉斯方程的解.

2. 证明，若 $z = f(x,y)$ 是一个锥面的方程，则

$$f_{xx}f_{yy} - f_{xy}^2 = 0.$$

3. 设 $f(x,y,z) = g(r)$, 其中 $r = \sqrt{x^2+y^2+z^2}$.

(a) 计算 $f_{xx}+f_{yy}+f_{zz}$.

(b) 证明，如果 $f_{xx}+f_{yy}+f_{zz}=0$, 则 $f(x,y,z)=\frac{a}{r}+b$, 其中 a 和 b 是常数.

4. 设 $f(x_1,x_2,\cdots,x_n)=g(r)$, 其中 $r=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$.

(a) 计算 $f_{x_1x_1}+f_{x_2x_2}+\cdots+f_{x_nx_n}$ (参看 1.4a, 练习 10).

(b) 解方程 $f_{x_1x_1}+f_{x_2x_2}+\cdots+f_{x_nx_n}=0$.

b. 例[1)]

1. 让我们考虑函数

$$u=\exp(x^2\sin^2 y+2xy\sin x\sin y+y^2).$$

我们令

$$u=e^{\xi+\eta+\zeta},\quad \xi=x^2\sin^2 y,\quad \eta=2xy\sin x\sin y,\quad \zeta=y^2,$$

便得到

$$\xi_x=2x\sin^2 y,\quad \eta_x=2y\sin x\sin y+2xy\cos x\sin y,\quad \zeta_x=0;$$

$$\xi_y=2x^2\sin y\cos y,\ \eta_y=2x\sin x\sin y+2xy\sin x\cos y,\ \zeta_y=2y;$$

$$u_\xi=u_\eta=u_\zeta=e^{\xi+\eta+\zeta}.$$

因此,

$$\begin{aligned}u_x=&2\exp(x^2\sin^2 y+2xy\sin x\sin y+y^2)(x\sin^2 y+y\sin x\sin y\\&+xy\cos x\sin y)\end{aligned}$$

和

$$\begin{aligned}u_y=&2\exp(x^2\sin^2 y+2xy\sin x\sin y+y^2)(x^2\sin y\cos y\\&+x\sin x\sin y+xy\sin x\cos y+y).\end{aligned}$$

---

1) 我们注意，下列的微分法也能直接完成，而不用关于多变量函数的链式法则.

2. 对函数

$$u = \sin(x^2 + y^2)$$

我们令 $\xi = x^2 + y^2$ 便得出

$$u_x = 2x\cos(x^2+y^2), \quad u_y = 2y\cos(x^2+y^2),$$
$$u_{xx} = -4x^2\sin(x^2+y^2) + 2\cos(x^2+y^2),$$
$$u_{xy} = -4xy\sin(x^2+y^2),$$
$$u_{yy} = -4y^2\sin(x^2+y^2) + 2\cos(x^2+y^2).$$

3. 对函数

$$u = \arctan(x^2 + xy + y^2),$$

作替换 $\xi = x^2, \eta = xy, \zeta = y^2$, 便导出

$$u_x = \frac{2x+y}{1+(x^2+xy+y^2)^2},$$
$$u_y = \frac{x+2y}{1+(x^2+xy+y^2)^2}.$$

c. 自变量的替换

锁链法则 (19) 应用到自变量替换的情形是特别重要的. 例如, 设 $u = f(\xi, \eta)$ 是两个自变量 ξ, η 的函数, 而 ξ, η 被看作 $\xi\eta$ 平面上的直角坐标. 我们可以在该平面上引进新的直角坐标 x, y (见第一卷第 405 页), 它们与 ξ, η 的关系由下式表示:

$$\xi = \alpha_1 x + \beta_1 y, \quad \eta = \alpha_2 x + \beta_2 y, \tag{24a}$$

或

$$x = \alpha_1 \xi + \alpha_2 \eta, \quad y = \beta_1 \xi + \beta_2 \eta. \tag{24b}$$

这里

$$\alpha_1 = \cos\gamma, \quad \alpha_2 = -\sin\gamma, \quad \beta_1 = \sin\gamma, \quad \beta_2 = \cos\gamma,$$

其中 γ 表示 ξ 轴的正向与 x 轴的正向之间的夹角．于是函数 $u=f(\xi,\eta)$ "变换" 成新的函数

$$u=f(\xi,\eta)=f(\alpha_1x+\beta_1y,\alpha_2x+\beta_2y)=F(x,y),$$

这是由 $f(\xi,\eta)$ 通过如第 59 页所述的复合步骤而形成的．我们说因变量 u 是 "相对于代替 ξ,η 的新的自变量 x,y 而说的"．

由第 61 页的微分法规则 (19) 立即得到

$$u_x=u_\xi\alpha_1+u_\eta\alpha_2,\quad u_y=u_\xi\beta_1+u_\eta\beta_2, \tag{25}$$

其中 u_x,u_y 表示函数 $u=F(x,y)$ 的偏导数，而 u_ξ,u_η 是函数 $f(\xi,\eta)$ 的偏导数．由此可见，当坐标轴旋转时，任何一个函数的偏导数按照与自变量的变换规则 (24b) 相同的规则而变换．这对将要讨论的空间坐标轴的旋转 (变换) 也是成立的 [1]．

另一个重要的自变量的替换是由直角坐标 (x,y) 变到**极坐标** (r,θ)．极坐标通过下述方程与直角坐标联系着：

$$x=r\cos\theta,\quad y=r\sin\theta, \tag{26a}$$

$$r=\sqrt{x^2+y^2},\ \theta=\arccos\frac{x}{\sqrt{x^2+y^2}}=\arcsin\frac{y}{\sqrt{x^2+y^2}}. \tag{26b}$$

将函数 $u=f(x,y)$ 用极坐标表示，我们有

$$u=f(x,y)=f(r\cos\theta,r\sin\theta)=F(r,\theta),$$

因而 u 就成为自变量 r 和 θ 的复合函数．从而，由锁链法则 (19) 我们得到

$$\begin{aligned}u_x&=u_rr_x+u_\theta\theta_x=u_r\frac{x}{r}-u_\theta\frac{y}{r^2}=u_r\cos\theta-u_\theta\frac{\sin\theta}{r},\\u_y&=u_rr_y+u_\theta\theta_y=u_r\frac{y}{r}+u_\theta\frac{x}{r^2}=u_r\sin\theta+u_\theta\frac{\cos\theta}{r}.\end{aligned} \tag{27}$$

---

1) 但是，一般说来，不适用于坐标变换的其他类型．

由此得出有用的方程

$$u_x^2 + u_y^2 = u_r^2 + \frac{1}{r^2}u_\theta^2. \tag{28}$$

由规则 (23a, b, c), 得出更高阶的导数

$$\begin{aligned}
u_{xx} =& u_{rr}\cos^2\theta + u_{\theta\theta}\frac{\sin^2\theta}{r^2} - 2u_{r\theta}\frac{\cos\theta\sin\theta}{r} \\
&+ u_r\frac{\sin^2\theta}{r} + 2u_\theta\frac{\cos\theta\sin\theta}{r^2}, \\
u_{xy} =& u_{yx} = u_{rr}\cos\theta\sin\theta - u_{\theta\theta}\frac{\cos\theta\sin\theta}{r^2} + u_{r\theta}\frac{\cos^2\theta - \sin^2\theta}{r} \\
&+ u_\theta\frac{\sin^2\theta - \cos^2\theta}{r^2} - u_r\frac{\sin\theta\cos\theta}{r}, \\
u_{yy} =& u_{rr}\sin^2\theta + u_{\theta\theta}\frac{\cos^2\theta}{r^2} + 2ur_\theta\frac{\cos\theta\sin\theta}{r} \\
&+ u_r\frac{\cos^2\theta}{r} - 2u_\theta\frac{\cos\theta\sin\theta}{r^2}.
\end{aligned}$$

由此可导出所谓拉普拉斯算式 Δu 的极坐标的表达式, 它出现在重要的 "拉普拉斯" 方程或 "位能" 方程 $\Delta u = 0$ (见第 36 页) 中:

$$\begin{aligned}
\Delta u = u_{xx} + u_{yy} &= u_{rr} + u_{\theta\theta}\frac{1}{r^2} + u_r\frac{1}{r} \\
&= \frac{1}{r^2}\left\{r\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{\partial^2 u}{\partial\theta^2}\right\}.
\end{aligned} \tag{29}$$

反过来, 我们可应用锁链法则把 u_r 和 u_θ 用 u_x 和 u_y 表示出来. 我们用这个方法求得

$$u_r = u_x x_r + u_y y_r = u_x\cos\theta + u_y\sin\theta, \tag{30a}$$

$$u_\theta = u_x x_\theta + u_y y_\theta = -u_x r\sin\theta + u_y r\cos\theta. \tag{30b}$$

我们也可通过对 u_r 和 u_θ 求解关系式 (27) 而得出这些方程. 顺便指出, 方程 (30a) 是早已出现过的, 即 u 沿向径 r 方向的导数的表达式, 见第 49 页.

一般地说, 只要给出了定义复合函数的关系式

$$u = f(\xi, \eta, \cdots),$$

$$\xi = \varphi(x,y), \quad \eta = \psi(x,y), \cdots,$$

我们就可把这些关系看作 u 是对新的自变量 x,y 而不是对 $\xi,\eta,\cdots$ 而说的. 自变量 x,y 和 $\xi,\eta,\cdots$ 的对应的值组, 都赋于 u 以同一个数值, 而无论 u 是作为 $\xi,\eta,\cdots$ 的函数 $f(\xi,\eta,\cdots)$ 还是作为 x,y 的函数

$$F(x,y) = f(\varphi(x,y),\psi(x,y),\cdots).$$

在一个复合函数 $u = f(\xi,\eta,\cdots)$ 的微分法中, 我们必须清楚地区分因变量 u 和函数 $f(\xi,\eta,\cdots)$, 后者根据自变量 $\xi,\eta,\cdots$ 的值给出 u 的值. 在 u 与自变量之间的函数关系确定之前, 微分法的记号 $u_\xi,u_\eta,\cdots$ 是没有意义的. 所以, 当处理复合函数

$$u = f(\xi,\eta,\cdots) = F(x,y)$$

时, 实际上不应该写 u_ξ,u_η 或 u_x,u_y 而应分别以 $f_\xi(\xi,\eta,\cdots)$, $f_\eta(\xi,\eta,\cdots)$ 或 $F_x(x,y)$, $F_y(x,y)$ 代替. 然而为了简洁起见, 在不致引起混淆的情况下, 简单的记号 u_ξ,u_η,u_x,u_y 仍是常用的. 于是锁链法则写成这样的形式

$$u_x = u_\xi\xi_x + u_\eta\eta_x, \quad u_y = u_\xi\xi_y + u_\eta\eta_y,$$

这使我们不必给出 u 与 ξ,η 或与 x,y 之间的函数关系的"名字" f 或 F.

下述例子说明这个事实, 量 u 关于一个给定的变量的导数依赖于 u 与**全部**自变量之间的函数关系的性质; 特别是它依赖于在微分过程中保持固定不变的那些自变量. 由"恒等变换" $\xi = x$, $\eta = y$, 函数 $u = 2\xi + \eta$ 变成 $u = 2x + y$, 我们有 $u_x = 2$, $u_y = 1$. 然而, 如果我们引进新的自变量 $\xi = x$ (与前面一样) 和 $\xi + \eta = v$, 就有 $u = x + v$, 因而 $u_x = 1$, $u_v = 1$. 这样, 由于另一个变量的不同的选择, 关于同一个自变量 x 的导数得出了不同的结果.

练 习 1.6 c

1. 设 $u=f(x,y)$, 其中 $x=r\cos\theta$, $y=r\sin\theta$. 试用 u_r 和 u_θ 来表示 $\sqrt{u_x^2+u_y^2}$.

2. 证明，在坐标系的旋转变换下关系式 $f_{xx}+f_{yy}$ 是不变的.

3. 证明，在线性变换 $x=\alpha\xi+\beta\eta$, $y=r\xi+\delta\eta$ 下，导数 $f_{xx}(x,y), f_{xy}(x,y), f_{yy}(x,y)$ 的变换规则分别与多项式

$$ax^2+2bxy+cy^2$$

的系数 a,b,c 的变换规则相同.

4. 给定 $z=r^2\cos\theta$, 这里 r 和 θ 是极坐标，求在点 $\theta=\frac{\pi}{4}$, $r=2$ 处的 z_x 和 z_y, 并用 z_x 和 z_y 表示 z_r 和 z_θ.

5. 由于变换 $\xi=a+\alpha x+\beta y$, $\eta=b-\beta x+\alpha y$, 其中 a,b,α,β 是常数而且 $\alpha^2+\beta^2=1$, 函数 $u(x,y)$ 变成 ξ,η 的函数 $U(\xi,\eta)$. 证明

$$U_{\xi\xi}U_{\eta\eta}-U_{\xi\eta}^2=u_{xx}u_{yy}-u_{xy}^2.$$

6. 说明当用变量 $z=x/\sqrt{y}$ 代替 y 时，关系式

$$T_y-T_{xx}$$

将如何变换.

7. (a) 证明，对于任意两次连续可微函数 f,g, 函数

$$h(x,y)=f(x-y)+g(x+y)$$

满足条件

$$h_{xx}=h_{yy}.$$

(b) 类似地，指出

$$H(x,y)=f(x-iy)+g(x+iy),$$

$i^2=-1$, 满足条件 $H_{xx}=-H_{yy}$.

问 题 1.6 c

1. 将拉普拉斯算式 $u_{xx}+u_{yy}+u_{zz}$ 转化成用三维极坐标 r,θ,φ 表示的形式，后者由下式定义：

$$x = r\sin\theta\cos\varphi,$$

$$y = r\sin\theta\sin\varphi,$$

$$z = r\cos\theta.$$

与 1.6a 问题 3 进行比较.

2. 寻求 a,b,c,d 的值，使在变换 $\xi = ax+by$, $\eta = cx+dy$ 下 (其中 $ad-bc\neq 0$) 方程 $Af_{xx}+2Bf_{xy}+Cf_{yy}=0$ 变成

(a) $f_{\xi\xi}+f_{\eta\eta}=0$,

(b) $f_{\xi\eta}=0$ (A,B,C 是常数).

这是否永远可能？

1.7 多元函数的中值定理与泰勒定理

a. 关于用多项式作近似的预备知识

在第一卷 (第五章第 506 页) 中我们看到，一个单变量的函数在一个给定点的邻域内能用一个 n 次多项式，即泰勒多项式近似代替，精确度比 n 阶更高，只要函数具有直至 $(n+1)$ 阶的导数. 用微分所表出的函数的线性部分作近似，只是实现这种较精确的近似式的第一步. 在多变量的情况，例如，两个自变量的函数的情况，我们也可在一个给定点的附近试图以一个 n 次多项式作为近似表达式. 或者说，我们希望用增量 h 和 k 的项组成的 "泰勒展式" 来逼近 $f(x+h,y+k)$.

通过一个简单的变换，这个问题可化为单变量的函数的情况. 代替刚才所考虑的 $f(x+h,y+k)$，我们引进一个附加变量 t 并把关系式

$$F(t) = f(x+ht, y+kt) \tag{31}$$

看成是 t 的函数，而把 x,y,h,k 暂时保持不变，当 t 在 0 与 1 之间变化时，以 $(x+ht,y+kt)$ 为坐标的点在连接 (x,y) 和 $(x+h,y+k)$ 的线段上变动. $F(t)$ 的以 t 的方幂组成的泰勒展式，当 $t=1$ 时将给出 $f(x+h,y+k)$ 的一个我们所希望的类型的近似表达式.

我们从计算 $F(t)$ 的导数开始. 我们假设下面将要写出的关于函数 $f(x,y)$ 的导数，都在一个完全包含 (x,y) 到 $(x+h,y+k)$ 这个线段的区域内是连续的，从锁链法则 (18) 立即得出 [1)]

$$F'(t)=hf_x+kf_y, \tag{32a}$$

$$F''(t)=h^2f_{xx}+2hkf_{xy}+k^2f_{yy}, \tag{32b}$$

$$\cdots\cdots\cdots$$

一般地说，由数学归纳法我们得出 n 阶导数的表达式

$$\begin{aligned}F^{(n)}(t)=&h^nf_{x^n}+\binom{n}{1}h^{n-1}kf_{x^{n-1}y}\\&+\binom{n}{2}h^{n-2}k^2f_{x^{n-2}y^2}+\cdots+k^nf_{y^n},\end{aligned} \tag{32c}$$

这里，像在第 56 页一样，能形象地写成

$$F^{(n)}(t)=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^nf$$

的形式. 对这个公式右边，符号的幂应先根据二项式定理展开，然后，将 $\partial/\partial x$, $\partial/\partial y$ 的幂乘 f'' 换成相应的 n 阶导数 $\partial^nf/\partial x^n$,

1) 从锁链法则我们有

$$F'(t)=\frac{d}{dt}f(x+ht,y+kt)=hf_\xi(\xi,\eta)+kf_\eta(\xi,\eta),$$

其中 $\xi=x+ht$, $\eta=y+kt$. 这里，我们将 $f_x(x+ht,y+kt)$ 作为 $f_\xi(x+ht,y+kt)$, 因为既然 x,y,h,k 被看作自变量，（仍由锁链法则）就有

$$\frac{\partial}{\partial x}f(x+ht,y+kt)=f_\xi(x+ht,y+kt).$$

$\partial^n f/\partial x^{n-1}\partial y, \cdots$. 在所有这些导数中，自变量 x 和 y 的位置应写入 $x+ht$ 和 $y+kt$.

练　习　1.7 a

1. 对 $F(t)=f(x+ht, y+kt)$ 求 $F'(1)$:

(a) $f(x,y)=\sin(x+y)$,

(b) $f(x,y)=\dfrac{y}{x}$,

(c) $f(x,y)=x^2+2xy^2-y^4$.

2. 求曲线 $z(t)=F(t)=f(x+ht, y+kt)$ 在 $t=1$ 处的斜率，其中 $x=0$, $y=1$, $h=\dfrac{1}{2}$, $k=\dfrac{1}{4}$, 并且

(a) $f(x,y)=x^2+y^2$,

(b) $f(x,y)=\exp[x^2+(y-1)^2]$,

(c) $f(x,y)=\cos\pi(y-1)\sin\pi x^2$.

b. 中值定理

在着手研究用多项式作更高阶的逼近以前，我们推导一个中值定理，它类似于对一元函数我们所早已知道的. 这个定理把差分 $f(x+h, y+k)-f(x,y)$ 与偏导数 f_x 和 f_y 联系了起来. 我们明确假设这些导数是连续的. 对函数 $F(t)$ 应用普通的中值定理，我们得到

$$\frac{F(t)-F(0)}{t}=F'(\theta t),$$

其中 θ 是介于 0 与 1 之间的一个数；利用 (31) 与 (32a) 即得

$$\frac{f(x+ht, y+kt)-f(x,y)}{t}$$
$$=hf_x(x+\theta ht, y+\theta kt)+kf_y(x+\theta ht, y+\theta kt).$$

令 $t=1$, 我们就得到想要求得的二元函数的中值定理，其形式为

$$f(x+h, y+k)-f(x,y)$$

$$= hf_x(x+\theta h, y+\theta k) + kf_y(x+\theta h, y+\theta k)$$
$$= hf_x(\xi,\eta) + kf_y(\xi,\eta). \tag{33}$$

因此，在两点 $(x+h, y+k)$ 和 (x,y) 处的函数值的差分等于在连接这两点的线段上的一个中间点 (ξ,η) 处的微分．值得注意的是，在 f_x 和 f_y 中都取同一个 θ 值．

正像对单变量函数（第一卷第 200 页）一样，中值定理能用于求得关于一个函数 $f(x,y)$ 的连续模，而且更确切地说，能证明上述函数 f 是利普希茨连续的．为了应用中值定理，我们必须用直线段连接两个点，使沿着这直线段 f 有定义．为此假设 $f(x,y)$ 的定义域 R 是凸的，亦即，连接 R 中任意两点的直线段完全落入 R 中．设 f 在 R 中是连续可微的，并设 M 是 f 的导数的绝对值的一个上界：

$$|f_x(x,y)| < M, \quad |f_y(x,y)| < M$$

对 R 中的一切 (x,y) 成立．于是，就能应用公式 (33) 并得出不等式

$$|f(x+h,y+k) - f(x,y)| \le |h||f_x(\xi,\eta)| + |k||f_y(\xi,\eta)|$$
$$\le |h|M + |k|M \le 2M\sqrt{h^2+k^2}. \tag{34}$$

因此 f 在距离为 $\rho = \sqrt{h^2+k^2}$ 的两点处的函数值的差不超过这个距离的一个固定的倍数（即 $2M\rho$）．这正是 f 的利普希茨连续性的意思．特别有

$$|f(x+h,y+k) - f(x,y)| < \varepsilon$$

对 $\sqrt{h^2+k^2} < \varepsilon/2M$．这样，$f$ 在 R 上是一致连续的，并且"连续模"$\delta = \varepsilon/2M$．

下述事实（其证明留给读者）是中值定理的一个简单的结果．

如果一个函数 $f(x,y)$ 的偏导数 f_x 和 f_y 存在而且在一个凸集上的每一点处都等于 0，那么这函数是一个常数．

练　习　1.7 b

1. 从几何上解释中值定理.

2. 对下列每一情况，求满足

$$hf_x(x+\theta h, y+\theta k)+kf_y(x+\theta h, y+\theta k)$$
$$=f(x+h, y+k)-f(x,y)$$

的 θ 值：

(a) $f(x,y)=xy+y^2$, $x=y=0$, $h=\frac{1}{2}$, $k=\frac{1}{4}$;

(b) $f(x,y)=\sin\pi(x+y)$, $x=y=\frac{1}{4}$, $h=\frac{1}{8}$, $k=\frac{1}{4}$.

3. 对函数

$$f(x,y)=\sin\pi x+\cos\pi y$$

用中值定理，来证明，存在一个数 θ $(0<\theta<1)$ 使

$$\frac{2}{\pi}=\cos\frac{\pi\theta}{2}+\sin\left[\frac{\pi}{2}(1-\theta)\right].$$

4. 对一个三元函数 $f(x,y,z)$ 推导中值定理.

5. 求一个数 θ $(0\le\theta\le1)$ 使

$$f\left(1,\frac{1}{2},\frac{1}{3}\right)=f_x\left(\theta,\frac{\theta}{2},\frac{\theta}{3}\right)+\frac{1}{2}f_y\left(\theta,\frac{\theta}{2},\frac{\theta}{3}\right)+\frac{1}{3}f_z\left(\theta,\frac{\theta}{2},\frac{\theta}{3}\right),$$

其中

(a) $f(x,y,z)=xyz$,

(b) $f(x,y,z)=x^2+y^2+2xz$.

问　题　1.7 b

1. 设 $f(x,y)$ 的定义域是一个多边形连通区域，亦即，假设这区域内的任意两点 P,Q 都能够用这区域内的一串线段 $\overline{P_0P_1}$, $\overline{P_1P_2}$, $\cdots$, $\overline{P_{n-1}P_n}$ 连接起来，其中 $P_0=P$, $P_n=Q$. 证明，若偏导数 f_x 和 f_y 在定义域内每一点处的值均为 0, 则 f 是常数.

c. 多个自变量的泰勒定理

我们对函数 $F(t)=f(x+ht,y+kt)$ 应用拉格朗日余项形式(参看第一卷第 499 页) 的泰勒公式，利用关系式 (32a, b, c) 求出 F 的导数，而后令 $t=1$, 我们就得到二元函数的**泰勒定理**

$$\begin{aligned}f(x+h,y+k)=&f(x,y)+\{hf_x(x,y)+kf_y(x,y)\}\\&+\frac{1}{2!}\{h^2f_{xx}(x,y)+2hkf_{xy}(x,y)+k^2f_{yy}(x,y)\}\\&+\cdots+\frac{1}{n!}\left\{h^nf_{x^n}(x,y)+\binom{n}{1}h^{n-1}kf_{x^{n-1}y}(x,y)\right.\\&\left.+\cdots+k^nf_{y^n}(x,y)\right\}+R_n,\end{aligned}\tag{35}$$

其中 R_n 表示余项

$$\begin{aligned}R_n=&\frac{1}{(n+1)!}\{h^{n+1}f_{x^{n+1}}(x+\theta h,y+\theta k)+\cdots\\&+k^{n+1}f_{y^{n+1}}(x+\theta h,y+\theta k)\},\end{aligned}\tag{36}$$

$0<\theta<1$. 这样，增量 $f(x+h,y+k)-f(x,y)$ 就被写成了 $1,2,\cdots,(n+1)$ 次齐次多项式的和，其中每个齐次式除去各自的因子

$$\frac{1}{1!},\frac{1}{2!},\cdots,\frac{1}{n!},\frac{1}{(n+1)!}$$

外，分别是 $f(x,y)$ 在点 (x,y) 处的第 1 阶，第 2 阶，$\cdots$，第 n 阶微分

$$\begin{aligned}df&=hf_x+kf_y=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)f,\\d^2f&=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^2f=h^2f_{xx}+2hkf_{xy}+k^2f_{yy},\\d^nf&=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^nf\\&=h^nf_{x^n}+\binom{n}{1}h^{n-1}kf_{x^{n-1}y}+\cdots+k^nf_{y^n},\end{aligned}$$

以及 f 在连接 (x,y) 和 $(x+h,y+k)$ 的线段上一个中间点处的 $(n+1)$ 阶微分 $d^{n+1}f$. 从而，泰勒定理可写成更简洁的形式：

$$f(x+h,y+k)=f(x,y)+df(x,y)+\frac{1}{2!}d^2f(x,y)+\cdots+\frac{1}{n!}d^nf(x,y)+R_n, \tag{37}$$

其中

$$R_n=\frac{1}{(n+1)!}d^{n+1}f(x+\theta h,y+\theta k),\quad 0<\theta<1. \tag{38}$$

在一般情况，余项 R_n 趋向于 0 比它的前项 d^nf 更高阶；亦即，当 $h\to 0$ 和 $k\to 0$ 时，有 $R_n=o\{\sqrt{(h^2+k^2)^n}\}$.

从对一元函数的泰勒定理通过 $(n\to\infty)$ 得出**无穷泰勒级数**，使我们导出了许多函数的幂级数展开式. 对多变量函数，这样的一种处理手法即使可能时，一般说来也是很麻烦的. 对我们来说，泰勒定理的重要性在于，将函数的增量 $f(x+h,y+k)-f(x,y)$ 分解成不同阶数的微分 $df,d^2f,\cdots$.

练　习　1.7 c

1. 求一个二次多项式，使它在原点的邻域内最接近于 $\sin x\sin y$.

2. 对 $f(x,y)=x^3+4y^2x$，求 $f\ (2.1,2.9)$ 的近似值.

3. 对 $f(x,y)=x/y+y/x$，估计用 $f(1,1)$ 近似代替 $f\ (0.9,0.9)$ 时的误差.

4. 把下列函数 $f(x+h,y+k)$ 展开成 h,k 的幂：

(a) $f(x,y)=x^3-2x^2y+y^2$,

(b) $f(x,y)=\cos(x+2y)$，在 $x=0,\ y=\dfrac{\pi}{2}$,

(c) $f(x,y)=x^4y+2y^2x-\sqrt{3x^2}$.

5. 将 $f(x,y,z)=xyz^2$ 展开成 $x,(y-1),(z+1)$ 的幂.

6. 求下列函数的泰勒展式的前几项，在原点 (0,0) 的一个邻域内：

(a) $z=\arctan\frac{y}{(x^2+1)}$,

(b) $z=\cosh x\sinh y$,

(c) $z=\cos x\cosh(x+y)$,

(d) $z=e^x\cos y$,

(e) $z=\frac{\sin x}{\cos y}$,

(f) $z=\log(1-x)\log(1-y)$,

(g) $z=e^{x^2-y^2}$,

(h) $z=\cos(x+y)e^{-x^2}$,

(i) $z=\cos(x\cos y)$,

(j) $z=\sin(x^2+y^2)$.

7. 估计用

$$1-\frac{1}{2}(x^2-y^2)$$

代替 $\cos x/\cos y$ 时的误差，对 $|x|,|y|<\frac{\pi}{6}$.

问　题　1.7 c

1. 求下列函数的泰勒级数并指出它们的有效区域：

(a) $\frac{1}{1-x-y}$,　　(b) e^{x+y}.

2. 证明：球面三角学中的余弦定律

$$\cos z=\cos x\cos y+\sin x\sin y\cos\theta$$

在原点的邻域内转化为欧几里得余弦定律

$$z^2=x^2+y^2-2xy\cos\theta.$$

3. 若 $f(x,y)$ 是一个连续函数且具有连续的一阶和二阶偏导数，则

$$f_{xx}(0,0)=\lim_{h\to+0}\frac{f(2h,e^{-1/2h})-2f(h,e^{-1/h})+f(0,0)}{h^2}.$$

4. 证明函数 $f(x,y)=\exp(-y^2+2xy)$ 能展成对 x 和 y 的一切值都收敛的级数

$$\sum_{n=0}^{\infty}\frac{H_n(x)}{n!}y^n,$$

其中多项式 $H_n(x)$, 即所谓哈密顿多项式, 满足

(a) $H_n(x)$ 是一个 n 次多项式,

(b) $H_n'(x)=2nH_{n-1}(x)$,

(c) $H_{n+1}-2xH_n+2nH_{n-1}=0$,

(d) $H_n''-2xH_n'+2nH_n=0$.

1.8 依赖于参量的函数的积分

多变量函数的重积分的概念将在第四章及第五章中才开始讨论. 目前我们仅考虑与这类函数有关的单重积分.

a. 例和定义

设 $f(x,y)$ 是 x 和 y 在矩形区域 $\alpha\le x\le\beta$, $a\le y\le b$ 上的连续函数, 我们可以把 x 看作是固定的而把 $f(x,y)$ 看成仅是 y 的函数并在区间 $a\le y\le b$ 上求积分. 这样我们得到表达式

$$\int_a^b f(x,y)dy,$$

它仍然依赖于量 x 的选取. 这样, 我们所考虑的就不只是一个积分而是一族积分 $\int_a^b f(x,y)dy$, 它由 x 取不同的值而得到. 量 x 在积分过程中保持不变, 而且可给它指定区间中任何一个值, 我们称它为一个参 (变) 量. 因而所说的这个通常的积分就成为参量 x 的一个函数.

作为参量的函数的积分在分析及其应用中是常见的. 例如, 作

变换 $xy = u$ 易得

$$\int_0^1 \frac{xdy}{\sqrt{1-x^2y^2}} = \arcsin x,$$

对 $-1 < x < 1$. 又如，我们可将指数函数看作一个参量的普通幂函数来求积分并写成

$$\int_0^1 y^x dy = \frac{1}{1+x},$$

其中假定 $x > -1$.

我们从几何上来阐明函数 $f(x,y)$ 的定义域，对参量 x 的固定值，作 y 轴的平行线，如图 1.15. 设 AB 是平行线与矩形相交的部分，被积函数 $f(x,y)$ 沿 AB 取值，因而它只是 y 的函数. 我们也可说**沿线段** AB 积分函数 $f(x,y)$.

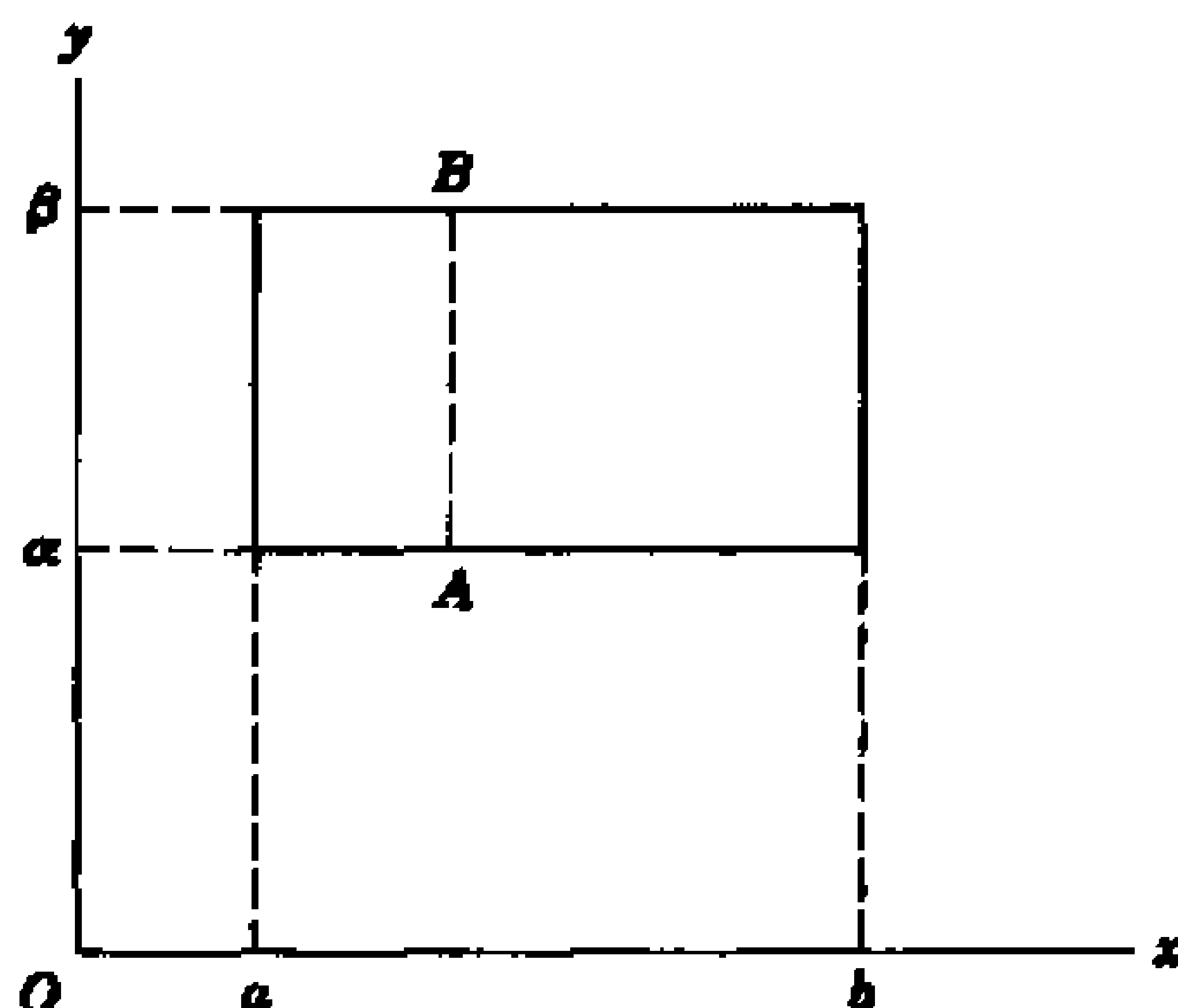


图 1.15

这种几何观点启示着一种推广. 若函数 $f(x,y)$ 的定义区域 R 取图 1.16 所示的形状，即任何一条平行于 y 轴的直线与边界最多交于两点，那么，对 x 的每一个固定的值，平行于 y 轴的直线与区域 R 相交的部分就是一个线段 AB (或一点 $A = B$)，因而我们仍可对函数 $f(x,y)$ 沿线段 AB 进行积分. 积分区间的始点和终点本身

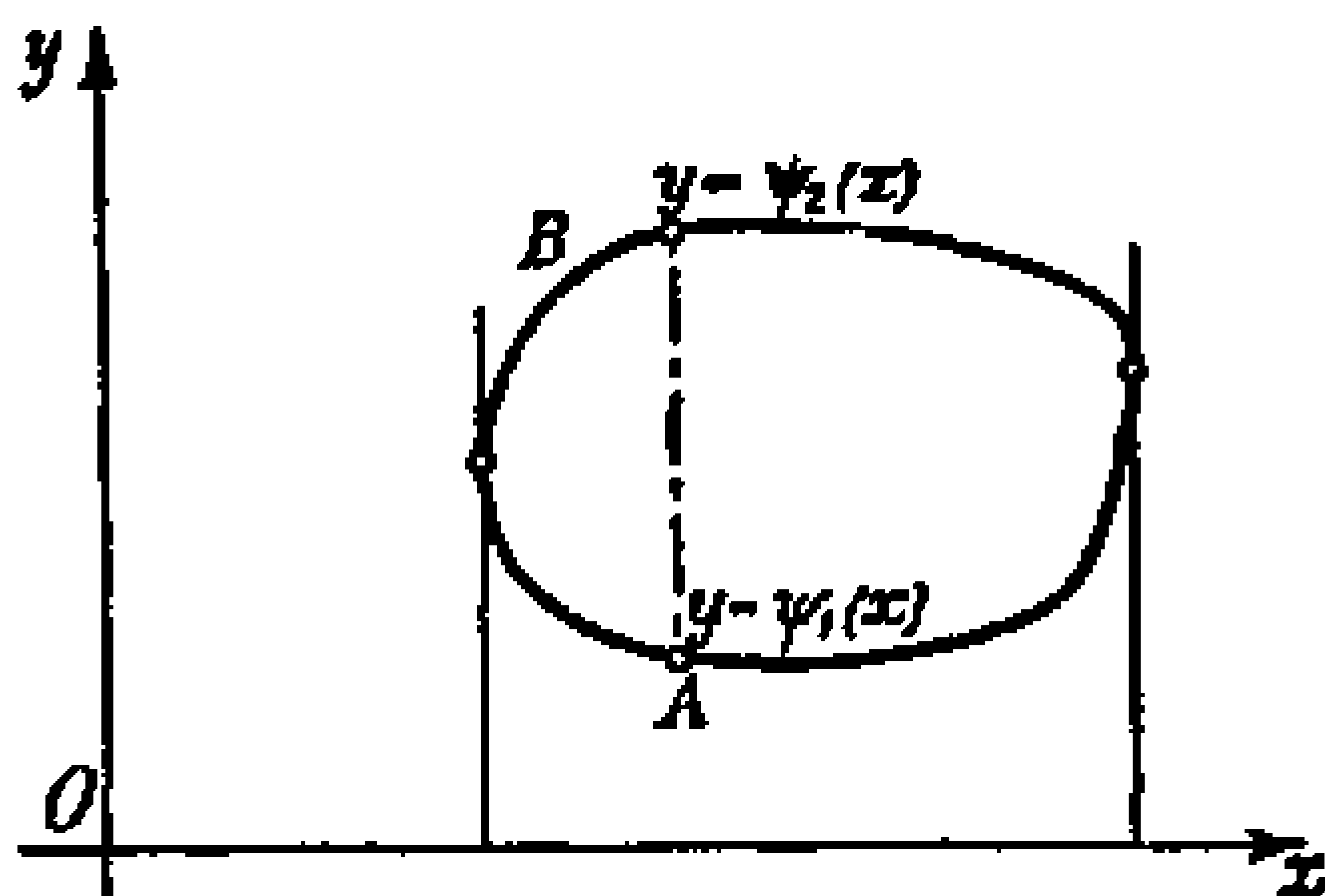


图 1.16

将随 x 而变化. 于是我们不得不考虑这种形式的积分:

$$\int_{\psi_1(x)}^{\psi_2(x)} f(x,y)dy = F(x), \tag{39}$$

在这种积分中 y 为积分变量, 而参量 x 在被积函数与积分限中都出现. 若我们用 x,y,z 空间的曲面 $z=f(x,y)$ 来表示函数 $f(x,y)$, 则对每一个正的函数 f, 我们可以考虑一个以 x,y 平面上的区域 R 为底, 以曲面 $z=f(x,y)$ 为顶, 以平行于 z 轴的直线为母线的柱面. x 的每一个固定值对应于一个平行于 y,z 平面的平面, 它与柱体相交成一个确定的平面区域. 这个平面区域的面积由积分 (39) 给出. 例如, 积分

$$\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1-x^2-y^2}dy$$

表示半球

$$0<z<\sqrt{1-x^2-y^2}$$

与平面 $x=$ 常数相交所得区域的面积.

b. 积分关于参量的连续性和可微性

如果 $f(x,y)$ **在闭矩形** $R:\alpha\le x\le\beta, a\le y\le b$ **上连续, 则积分**

$$F(x)=\int_a^b f(x,y)dy$$

是参量 x 在区间 $\alpha \le x \le \beta$ 上的连续函数.

因为

$$|F(x+h)-F(x)| = \left|\int_a^b (f(x+h,y)-f(x,y))dy\right|$$
$$\le \int_a^b |f(x+h,y)-f(x,y)|dy.$$

由于 $f(x,y)$ 的一致连续性，对充分小的 h 的值，右边的被积函数作为 y 的函数可以一致地小到我们所希望的程度，由此即得以上结论.

其次我们考虑微分 $F(x)$ 的可能性. 首先考虑积分限固定且函数 $f(x,y)$ 在闭矩形 R 中有连续偏导数 f_x 的情况[1]. 我们将证明，先对 y 求积分然后对 x 求微分的运算，可以用交换这两步骤次序的运算来代替.

定理. 若函数 $f(x,y)$ 在闭矩形 $\alpha \le x \le \beta$, $a \le y \le b$ 中连续且有关于 x 的连续导数，则可以在积分号下关于参量微分这个积分，即

$$\frac{d}{dx}F(x) = \frac{d}{dx}\int_a^b f(x,y)dy = \int_a^b f_x(x,y)dy. \tag{40}$$

而且，$F'(x)$ 是 x 的一个连续函数.

在证明定理之前，我们指出，这个定理给出了下述事实 (在第 38 页中已建立的) 的简单证明，在求一个函数 $g(x,y)$ 的混合导数 g_{xy} 时，只要 g_y 和 g_{xy} 都连续且 g_x 存在，则微分的次序可以交换. 因为若令 $f(x,y) = g_y(x,y)$，我们有

$$g(x,y) = g(x,a) + \int_a^y f(x,\eta)d\eta.$$

因 $f(x,y)$ 在矩形 $\alpha \le x \le \beta$, $a \le y \le b$ 中有关于 x 的连续微商，立即有

$$g_x(x,y) = g_x(x,a) + \int_a^y f_x(x,\eta)d\eta,$$

---

1) 这意味着 f_x 在开矩形上存在而且能扩充为闭矩形上的连续函数 (见第 46 页).

因而根据微积分基本定理便有

$$g_{yx}(x,y) = f_x(x,y).$$

又因根据 f 的定义有 $f_x(x,y) = g_{xy}(x,y)$, 由此看出 $g_{yx} = g_{xy}$.

定理的证明. 若 x 与 $x+h$ 都属于区间 $\alpha \leq x \leq \beta$, 就能写成

$$\begin{aligned} F(x+h) - F(x) &= \int_a^b f(x+h,y)dy - \int_a^b f(x,y)dy \\ &= \int_a^b [f(x+h,y) - f(x,y)]dy. \end{aligned}$$

由假设 $f(x,y)$ 关于 x 可微，根据微积分学中的微分中值定理的通常形式便得到 [1]

$$f(x+h,y) - f(x,y) = hf_x(x+\theta h,y), \quad 0 < \theta < 1.$$

且由于假设了导函数 f_x 在闭矩形上连续因而一致连续，差分

$$f_x(x+\theta h,y) - f_x(x,y)$$

的绝对值可小于任何正数 ε, 对一切满足 $|h| < \delta$ 的 h 都成立，此处 $\delta = \delta(\varepsilon)$ 不依赖于 x 和 y. 于是

$$\begin{aligned} &\left|\frac{F(x+h) - F(x)}{h} - \int_a^b f_x(x,y)dx\right| \\ =&\left|\int_a^b f_x(x+\theta h,y)dy - \int_a^b f_x(x,y)dy\right| \\ \leq &\int_a^b \varepsilon dy = \varepsilon(b-a). \end{aligned}$$

对 $|h| < \delta(\varepsilon)$ 成立，只要 $h \neq 0$. 这意味着，关系式

$$\lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \int_a^b f_x(x,y)dy = F'(x)$$

1) 此处 θ 依赖于 y 而且甚至可以不连续地随 y 变化. 这是无关紧要的，因为由等式 $f_x(x+\theta h,y) = h^{-1}[f(x+h,y) - f(x,y)]$ 可立即看出 $f_x(x+\theta h,y)$ 是 x 和 y 的连续函数因而是可积的.

成立. 这证明了 $F'(x)$ 的存在和公式 (40). $F'(x)$ 的连续性可根据被积函数 $f_x(x,y)$ 的连续性立即推出 (见第 82 页).

当**积分限**中包含参量时, 用类似的方法可以建立积分的连续性以及积分关于参量的微分法则.

例如, 若我们要想微分

$$F(x)=\int_{\psi_1(x)}^{\psi_2(x)} f(x,y)dy,$$

我们先考虑关系式

$$F(x)=\int_u^v f(x,y)dy=\varphi(u,v,x),$$

其中 $u=\psi_1(x)$, $v=\psi_2(x)$. 我们假设 $\psi_1(x)$ 和 $\psi_2(x)$ 在区间 $\alpha\le x\le\beta$ 上有连续的一阶导数且

$$a<\psi_1(x)<\psi_2(x)<b$$

对 $\alpha<x<\beta$ 成立. 又设 $f(x,y)$ 和 $f_x(x,y)$ 在集合 $\alpha\le x\le\beta$, $a\le y\le b$ 上是连续的. 这样, 对

$$\alpha\le x\le\beta,\quad a\le u\le b,\quad a\le v\le b$$

就定义了三个自变量 u,v,x 的一个函数 ϕ. 而且这函数有连续的偏导数, 因为由公式 (40), 有

$$\phi_x(u,v,x)=\frac{\partial}{\partial x}\int_u^v f(x,y)dy=\int_u^v f_x(x,y)dy,$$

而根据微积分基本定理 (第一卷第 207 页), 有

$$\begin{aligned}\phi_v(u,v,x)&=\frac{\partial}{\partial v}\int_u^v f(x,y)dy=f(x,v),\\ \phi_u(u,v,x)&=\frac{\partial}{\partial u}\int_u^v f(x,y)dy=-\frac{\partial}{\partial u}\int_v^u f(x,y)dy\\ &=-f(x,u).\end{aligned}$$

对复合函数 $F(x)=\phi(\psi_1(x),\psi_2(x),x)$ 应用第 67 页上的微分锁链法则 (18) 便得到

$$F'(x)=\varphi_u\psi_1'(x)+\varphi_v\psi_2'(x)+\varphi_x,$$

这证明了 $F(x)$ 的连续导数存在 $(\alpha<x<\beta)$, 并得到公式

$$\begin{aligned}\frac{d}{dx}\int_{\psi_1(x)}^{\psi_2(x)}f(x,y)dy=&\int_{\psi_1(x)}^{\psi_2(x)}f_x(x,y)dy-\psi_1'(x)f(x,\psi_1(x))\\&+\psi_2'(x)f(x,\psi_2(x)).\end{aligned}\tag{41}$$

作为例子, 对函数

$$F(x)=\int_0^x\sin(xy)dy,$$

我们得到

$$\frac{dF(x)}{dx}=\int_0^x y\cos(xy)dy+\sin(x^2).$$

对于

$$F(x)=\int_0^1\frac{xdy}{\sqrt{1-x^2y^2}}=\arcsin x,$$

当 $-1<x<+1$, 我们得到关系式

$$F'(x)=\int_0^1\frac{dy}{\sqrt{(1-x^2y^2)^3}}=\frac{1}{\sqrt{1-x^2}},$$

这结果读者可以直接验证.

再看一序列积分

$$F_n(x)=\int_0^x\frac{(x-y)^n}{n!}f(y)dy,\quad F_0(x)=\int_0^x f(y)dy,\tag{42}$$

其中 n 是任意正整数且 $f(y)$ 仅是 y 在所考虑的区间上的连续函数. 由于对上限 x 求导的结果为 0, 由规则 (41) 得递归公式

$$F_n'(x)=F_{n-1}(x)\quad(n=1,2,3,\cdots).$$

从 $F_0'(x) = f(x)$ 立即得出

$$F_n^{(n+1)}(x) = f(x). \tag{42a}$$

因而 $F_n(x)$ 是这样的函数：它的 $(n+1)$ 阶导数等于 $f(x)$，它本身及其前 n 阶导数在 $x=0$ 处取零值；它是由 $F_{n-1}(x)$ 从 0 到 x 积分而得. 因此 $F_n(x)$ 是以 0 与 x 为积分限由 $f(x)$ 经 $(n+1)$ 次积分而得的函数：

$$\begin{aligned} &F_0(x) = \int_0^x f(y)dy, \quad F_1(x) = \int_0^x F_0(y)dy, \\ &F_2(x) = \int_0^x F_1(y)dy, \cdots, F_n(x) = \int_0^x F_{n-1}(y)dy. \end{aligned} \tag{42b}$$

这种重复积分能用函数 $\frac{(x-y)^n}{n!}f(y)$ 关于 y 的一个简单积分所代替.

关于参量求微分的上述这些法则，当由积分号下求微分而得的函数不处处连续时，往往也仍然有效. 对这些情况，代替应用一般的准则，对每一个特殊的情况，直接验证上述微分是否可行反而更方便些.

作为例子，我们考虑椭圆积分 (见第一卷第 342 页).

$$F(k) = \int_{-1}^{+1} \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}; \quad k^2 < 1.$$

函数

$$f(k,x) = \frac{1}{\sqrt{(1-x^2)(1-k^2x^2)}}$$

在 $x=+1$ 和 $x=-1$ 处是不连续的，但积分 (作为广义积分) 是有意义的. 作关于参数 k 的形式微分得

$$F'(k) = \int_{-1}^{+1} \frac{kx^2dx}{\sqrt{(1-x^2)(1-k^2x^2)^3}}.$$

为了考察这等式是否成立，我们重复推导微分公式的论据. 这样就有

$$\frac{F(k+h)-F(k)}{h}=\int_{-1}^{+1} f_k(k+\theta h,x)dx$$
$$=\int_{-1}^{+1}\frac{(k+\theta h)x^2dx}{\sqrt{(1-x^2)[1-(k+\theta h)^2x^2]^3}}.$$

这个表达式与上面通过形式微分所得的积分之间的差是

$$\Delta=\int_{-1}^{+1}\frac{x^2}{\sqrt{1-x^2}}\left(\frac{k+\theta h}{\sqrt{[1-(k+\theta h)^2x^2]^3}}-\frac{k}{\sqrt{(1-k^2x^2)^3}}\right)dx.$$

我们必须证明这个积分随 h 趋于 0. 为此目的，我们在 k 周围划定一个不包含数值 ± 1 的区间 $k_0\le k\le k_1$, 并且选取 h 足够小使 $k+\theta h$ 落在此区间内. 函数

$$\frac{k}{\sqrt{(1-k^2x^2)^3}}$$

在闭区域 $-1\le x\le 1, k_0\le k\le k_1$ 上是连续的，因而一致连续. 从而差

$$\left|\frac{k+\theta h}{\sqrt{[1-(k+\theta h)^2x^2]^3}}-\frac{k}{\sqrt{(1-k^2x^2)^3}}\right|$$

小于一个不依赖于 x 和 k 并且随 h 趋于 0 的数 ϵ. 从而

$$|\Delta|\le\int_{-1}^{+1}\frac{x^2dx}{\sqrt{1-x^2}}\cdot\epsilon=M\epsilon,$$

其中 M 是一个不依赖于 ϵ 的常数. 这就是说，积分 Δ 随 h 趋于零而趋于零. 这就是我们所要证明的.

因而在这种情况下，积分号下求微分仍然是可行的. 类似的考虑可应用于其他的情况.

具有无限的积分区间且依赖于参量的广义积分将在第四章中讨论.

练 习 1.8 b

1. 设

$$F(k)=\int_a^b \alpha(x)\beta(x,k)dx,$$

其中 $\beta(x,k)$ 和 $\beta_k(x,k)$ 对 $a\le x\le b$, $k_0<k<k_1$ 是连续的，又 $\alpha(x)$ 对 $a<x<b$ 连续，而且 $\int_a^b |\alpha(x)|$ 作为广义积分存在. 求证

$$F'(k)=\int_a^b \alpha(x)\beta_k(x,k)dx,\quad k_0<k<k_1.$$

2. 设

$$F(k)=\int_0^1 (x-1)x^k\log^{-1}x dx,\quad -1<k.$$

证明

(a) $\lim\limits_{k\to\infty} kF(k)=1$,

(b) $F(k)=\log\dfrac{2+k}{1+k}$.

c. 积分 (次序) 的互换. 函数的光滑化

第 82 页上关于积分号下求微分的定理导致一个重要的结果，即可以交换积分的次序.

设 $f(x,y)$ 在由

$$a\le x\le b,\quad \alpha\le y\le\beta \tag{42c}$$

所确定的矩形 R 上连续. 则积分

$$I=\int_a^b d\xi\int_\alpha^\beta f(\xi,\eta)d\eta \text{ 和 } J=\int_\alpha^\beta d\eta\int_a^b f(\xi,\eta)d\xi \tag{42d}$$

具有相同的值. 我们称这个值为 f 在矩形(42c)上的二重积分.

作为例子，我们考虑矩形 $0 \le x \le 1, 0 \le y \le \frac{\pi}{2}$ 上的函数

$$f(x,y) = y\sin(xy).$$

这里

$$I = \int_0^1 d\xi \int_0^{\pi/2} \eta \sin(\xi\eta) d\eta = \int_0^1 \left(-\frac{\pi\cos(\pi\xi/2)}{2\xi} + \frac{\sin(\pi\xi/2)}{\xi^2}\right) d\xi$$
$$= \frac{\pi}{2} - 1,$$

$$J = \int_0^{\pi/2} d\eta \int_0^1 \eta \sin(\xi\eta) d\xi = \int_0^{\pi/2} (1 - \cos\eta) d\eta = \frac{\pi}{2} - 1.$$

为一般地证明等式 $I = J$，我们引进两个变上限的积分

$$v(x,y) = \int_\alpha^y f(x,\eta) d\eta, \quad u(x,y) = \int_a^x v(\xi,y) d\xi.$$

应用公式 (40) 有

$$u_y(x,y) = \int_a^x v_y(\xi,y) d\xi = \int_a^x f(\xi,y) d\xi,$$

从而

$$u(x,y) = u(x,\alpha) + \int_\alpha^y u_y(x,\eta) d\eta = \int_\alpha^y d\eta \int_a^x f(\xi,\eta) d\xi.$$

以 $x = b, y = \beta$ 代入即得 $I = J$.

我们已建立了函数 $u(x,y)$ 与一个在矩形 R 上连续的函数 $f(x,y)$ 间的联系： $u(x,y)$ 有连续的一阶偏导数

$$u_x(x,y) = \int_\alpha^y f(x,\eta) d\eta, \quad u_y(x,y) = \int_a^x f(\xi,y) d\xi,$$

还有一个连续的混合二阶偏导数

$$u_{xy}(x,y) = f(x,y).$$

为了"光滑化"f，即为了构造一个一致逼近于 f 而有连续偏导数的函数，我们将利用这样的函数 u.

在应用技巧上，往往需要用一个相近的光滑函数来代替一个连续函数 f (它本身往往只是某个不完全已知的物理量的近似量). 由魏尔斯特拉斯逼近定理 (第一卷第 686 页) 我们知道，在一个区间上连续的单变量的函数可以用具有一切阶导数的多项式一致逼近. 对在一个矩形上连续的二元函数 $f(x,y)$, 类似的定理仍然成立.

我们能借助 "平均" 函数 $f(x,y)$ 的方法构造一个具有适当光滑次数的简单逼近. 一个方便的办法是扩充 f 的定义，从矩形区域 (42c) 到整个 x,y 平面，使 f 处处连续[1]. 对任意 $h>0$, 我们在以 (x,y) 为中心以 $2h$ 为边长且四边平行于坐标轴的正方形上作 f 的均值：

$$
\begin{aligned}
F_h(x,y) &= \frac{1}{4h^2}\int_{x-h}^{x+h} d\xi \int_{y-h}^{y+h} f(\xi,\eta)d\eta \\
&= [u(x+h,y+h)-u(x+h,y-h) \\
&\quad - u(x-h,y+h)+u(x-h,y-h)]/4h^2. \qquad (42e)
\end{aligned}
$$

显然 $F_h(x,y)$ 有连续的一阶偏导数和连续的混合二阶偏导数[2]. 为要说明对小的 h, $F_h(x,y)$ 逼近于 $f(x,y)$, 我们注意

$$
\begin{aligned}
&F_h(x,y)-f(x,y) \\
&= \frac{1}{4h^2}\int_{x-h}^{x+h} d\xi \int_{y-h}^{y+h} [f(\xi,\eta)-f(x,y)]d\eta. \qquad (42f)
\end{aligned}
$$

由 f 在某个其内部包含 R 的矩形 R' 内是一致连续的, 我们知道, 对给定的 ϵ 和充分小的 h, 在每一个包含在 R' 内的边长为 $2h$ 的正方形内，f 的变化将小于 ϵ. 于是在 (42f) 中我们有 $|f(\xi,\eta)-f(x,y)|<\epsilon$ 以及 $|F_h(x,y)-f(x,y)|<\epsilon$. 因而 $\lim\limits_{h\to 0} F_h(x,y)=f(x,h)$ 对 R 中的点 (x,y) 一致成立. 这样，我们就找到了一个任意接近于 $f(x,y)$ 的光滑函数 $F_h(x,y)$.

---

1) 这是能达到的，可这样来延续 f: 当沿着垂直于矩形的四边的射线，f 取常数，然后对平面上剩余的点，再这样延续：当沿着从矩形的四个角出发的射线，f 也取常数.

2) 为使 $F_h(x,y)$ 对矩形 R 中的一切点有定义，我们可使 f 定义在稍微超出 R 的区域上.

1.9 微分与线积分

a. 线性微分型

我们在 1.5d 中定义了函数 $u=f(x,y,z)$ 的全微分 du, 其表达式是

$$du=\frac{\partial f(x,y,z)}{\partial x}dx+\frac{\partial f(x,y,z)}{\partial y}dy+\frac{\partial f(x,y,z)}{\partial z}dz. \tag{43}$$

多元函数微分的这个定义是得到了 **微商的锁链法则** 的启发的. 因为如果 x,y,z 是变量 t 的函数

$$x=\varphi(t),\quad y=\psi(t),\quad z=\chi(t), \tag{44}$$

那么复合函数 $u=f[\varphi(t),\psi(t),\chi(t)]$ 的微商的锁链法则 (19) 就是

$$\frac{du}{dt}=\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}+\frac{\partial f}{\partial z}\frac{dz}{dt}. \tag{45}$$

而单个变量 t 的函数 u 的微分则是定义为 $du=\dfrac{du}{dt}dt$ 的; 所以由 (45) 就有

$$\begin{aligned}du&=\left(\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}+\frac{\partial f}{\partial z}\frac{dz}{dt}\right)dt\\&=\frac{\partial f}{\partial x}\frac{dx}{dt}dt+\frac{\partial f}{\partial y}\frac{dy}{dt}dt+\frac{\partial f}{\partial z}\frac{dz}{dt}dt,\end{aligned}$$

只要我们还记得 x,y,z (作为 t 的函数) 的微分是

$$dx=\frac{dx}{dt}dt,\quad dy=\frac{dy}{dt}dt,\quad dz=\frac{dz}{dt}dt,$$

那么上式就与 (43) 形式上完全一致了. 于是从 (43) 式所定义的微分 $du=df(x,y,z)$ 可以立刻得到 u"沿着任一曲线"(其参数方程为 (44)) 的微分 $du=\dfrac{du}{dt}dt$.

由 (43) 式所定义的微分 du 是六个变量 x,y,z,dx,dy,dz 的函数，它对变量 dx,dy,dz 是线性、齐次的[1]，系数是 x,y,z 的函数．（当然，就像在第 51 页中所解释的，并非任何时候都要求微分 dx,dy,dz"很小"；只有当我们要用 du 去作增量

$$\Delta u = f(x+dx,y+dy,z+dz) - f(x,y,z)$$

的一个近似值时才提出这个限制．)

在 x,y,z 空间中最一般的线性微分型的表达式是

$$L = A(x,y,z)dx + B(x,y,z)dy + C(x,y,z)dz. \tag{46}$$

这是六个变量 x,y,z,dx,dy,dz 的一个函数，对"微分"变量 dx,dy,dz 是线性型，系数依赖于 x,y,z．函数的全微分 du 是特殊的线性微分型 L，它的系数为：

$$A = \frac{\partial f(x,y,z)}{\partial x}, \quad B = \frac{\partial f(x,y,z)}{\partial y}, \quad C = \frac{\partial f(x,y,z)}{\partial z}, \tag{47}$$

其中 $f = f(x,y,z)$ 是一个适当的函数．如果一个微分型 L 是某个函数的全微分，我们就说它是一个**恰当**微分型或者说它是**可积的**．并不是每一个微分型都是可积的；若要 L 可积，它的系数 A,B,C 必须满足某种"可积条件"：

若线性微分型 L **的系数** A,B,C **属于** C^1 **类(就是说有连续的一阶微商**，见第 46 页)，**又如果** L **是恰当的，则下列方程成立**：

$$\frac{\partial B}{\partial z} - \frac{\partial C}{\partial y} = 0, \quad \frac{\partial C}{\partial x} - \frac{\partial A}{\partial z} = 0, \quad \frac{\partial A}{\partial y} - \frac{\partial B}{\partial x} = 0. \tag{48}$$

方程 (48) 是二阶微商可交换性法则的简单推论．因为如果 A,B,C 有连续的一阶微商又可以写成 (47) 的形式，那么 f 就有连续的二

1) 三个变量 ξ,η,ζ 的线性函数的普遍形式是 $A\xi + B\eta + C\zeta + D$，其中系数 A,B,C,D 不依赖于 ξ,η,ζ；当 $D = 0$ 时，称这线性函数为"齐次的"，或者说它是一个"线性型"(见第 14 页)．

阶微商. 根据第 37 页的定理混合微商与微分次序无关. 这样一来, 就有

$$\frac{\partial A}{\partial y}=\frac{\partial}{\partial y}\frac{\partial f}{\partial x}=\frac{\partial}{\partial x}\frac{\partial f}{\partial y}=\frac{\partial B}{\partial x},$$

类似地可得 (48) 中另外两个等式.

作为例子, 线性微分型

$$L=ydx+zdy+xdz$$

不是可积的, 因为

$$\frac{\partial B}{\partial z}-\frac{\partial C}{\partial y}=\frac{\partial z}{\partial z}-\frac{\partial x}{\partial y}=1\neq 0.$$

而线性微分型

$$L=yzdx+zxdy+xydz$$

满足可积条件 (48). 事实上它是函数 $u=xyz$ 的全微分 du. 至于在什么范围里条件 (48) 对于 L 是全微分也是**充分**的, 我们将在第 1.10 节中讨论.

对三维以外其他维数可以得到类似的可积条件. 对于两个独立变量 x,y 而言, 一般的线性微分型是 $L=A(x,y)dx+B(x,y)dy$. 如果 L 是一个函数 $u=f(x,y)$ 的全微分 du, 那么系数 A,B 应该满足方程

$$\frac{\partial A}{\partial y}-\frac{\partial B}{\partial x}=0.$$

另一方面, 在四维的情形下, 对应于方程 (48) 我们得到**六个**可积条件, 它们是由四元函数 f 的所有可能的混合二阶微商构成的.

考虑微分型, 甚至不是恰当微分的微分型有什么意义呢? 其理由为: 沿着参量方程

$$x=\varphi(t),\quad y=\psi(t),\quad z=\chi(t)$$

给出的曲线 C, L 就成为一个一元函数的微分

$$L=\left(A\frac{dx}{dt}+B\frac{dy}{dt}+C\frac{dz}{dt}\right)dt.$$

这个函数就是由不定积分

$$\int L = \int \left(A\frac{dx}{dt} + B\frac{dy}{dt} + C\frac{dz}{dt} \right) dt$$

给出的.

b. 线性微分型的线积分

为了讨论线性微分型在曲线上的积分，对有向弧和闭曲线的概念和性质有一个清楚的形象是很重要的. 望读者复习第一卷第 376–386 页，那里一切有关的论述是对**平面曲线**而作的，但可以同样地应用到任何维数的空间曲线上[1]. 不失一般性，我们来讨论 x, y, z 三维空间中的曲线上的积分.

一个**简单弧** Γ 是点 $P = (x, y, z)$ 的一个集合，这些点可以用参量形式表示出来：

$$x = \varphi(t), \quad y = \psi(t), \quad z = \chi(t); \quad a \leq t \leq b, \tag{49}$$

其中 φ, ψ, χ 是 t 在 $a \leq t \leq b$ 上的连续函数，并且这区间内不同的 t 对应到不同的点 P. 参量式 (49) 构造了一个一对一的连续映像把 t 轴上的一个区间映到空间的集合 Γ 上[2]. 同一个简单弧 Γ 可以有许多不同的参量表示. 最一般的一个可以从特殊的表示式 (49) 得到，这只需取任意一个把区间 $\alpha \leq \tau \leq \beta$ 映到区间 $a \leq t \leq b$ 上的单调连续函数 $\mu(\tau)$，再令

$$x = \psi[\mu(\tau)],\ y = \psi[\mu(\tau)],\ z = \chi[\mu(\tau)],\ \alpha \leq \tau \leq \beta. \tag{50}$$

---

1) 曲线的“正和负侧”的概念和“顺时针逆时针方向”的概念是二维特有的.

2) 由于假设了函数 φ, ψ, χ 的连续性，所以从 t 到 P 的映像的连续性是显然的. 重要的是来验证逆映像 $P \to t$ 也是连续的. 这意思是：给定了 Γ 上一个收敛到点 P 的点列 P_n，则对应的参量值 t_n 就收敛到 P 的参量值. 为了证明这点，我们注意到根据**有界闭区间的列紧性**(第一卷第 105 页) 有一个 t_n 的子序列收敛到某个值 t, $a \leq t \leq b$. 由原映像的连续性，t 被映到 P_n 的极限 P 上. 因为假设了映像有 1-1 的特性，t 是被 P 唯一地确定的. 所以 t_n 的每一个收敛子序列都有相同的极限，就是 P 所对应的参量值. 这就证明了整个序列 t_n 收敛到 t.

对于 Γ 的任何一个参量表示式 (49), 我们有两种方法排列 Γ 上的点的顺序, 它们对应于按照 t 的增加或减少的顺序. 选择了两个顺序之一, 就使简单弧 Γ 成为 **有向简单弧** Γ^*. 如果 Γ^* 的方向对应于 t 增加, 我们就说 Γ^* 的方向关于参量 t 是 **正的**; 如果 Γ^* 的方向对应于 t 减少, 就说 Γ^* 的方向是 **负的**. 若有向简单弧与 Γ^* 有相反的方向, 就记为 $-\Gamma^*$. 如果我们知道了 Γ 上任何两个点 P_0, P_1 的顺序, 那么 Γ 的方向就完全确定了. 如果 Γ^* 关于参数 t 是正向的, 又如果 t_0 和 t_1 分别是点 P_0 和 P_1 的参量值, 那么 $t_0 < t_1$ 就意味着在 Γ^* 上 P_1 **跟随** P_0 或 P_0 **先于** P_1 (图 1.17).

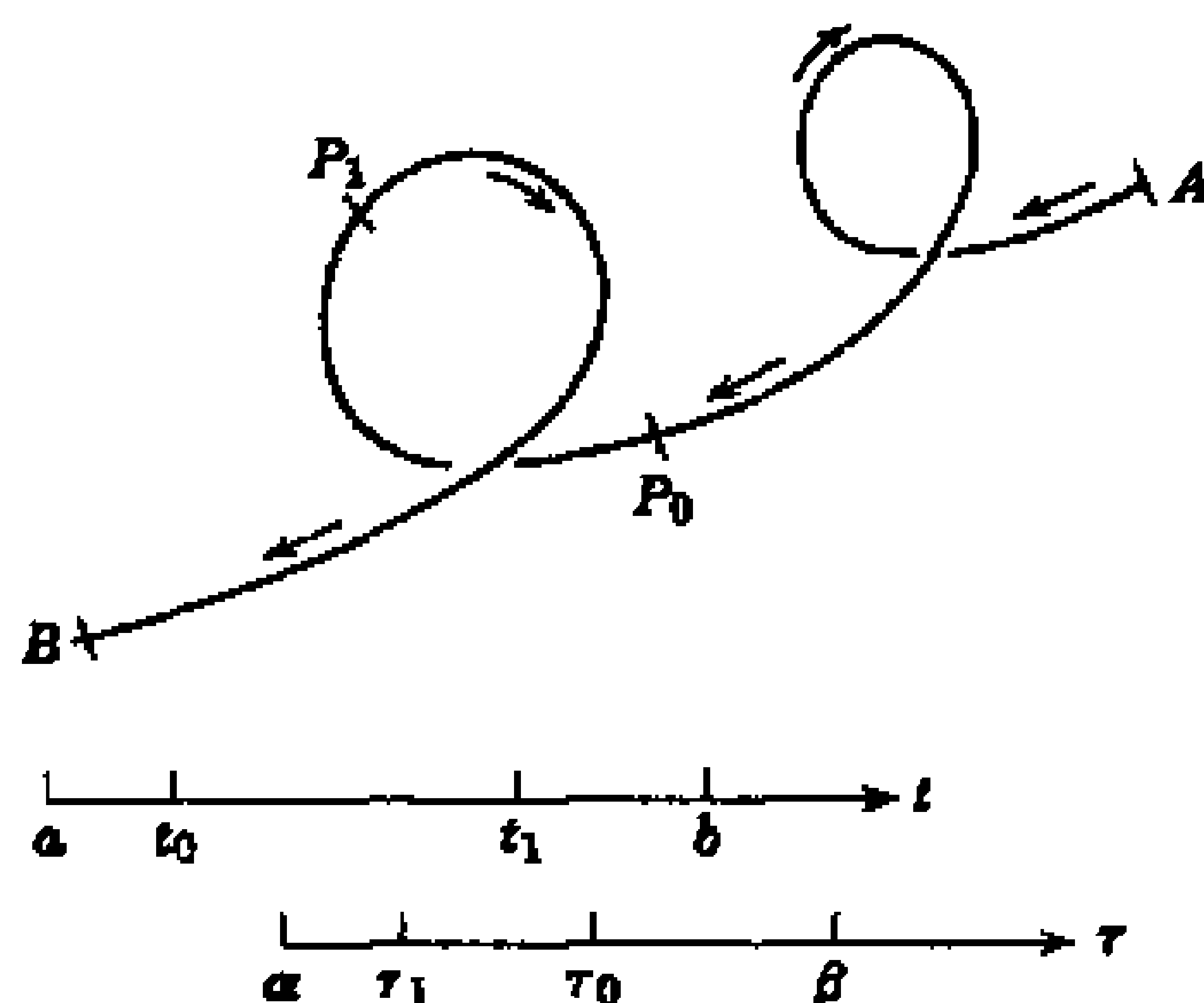


图 1.17 空间简单弧关于参量 τ 定向是负的, 关于参量 $t = \mu(\tau)$ 定向是正的, 这里 $\mu(\alpha) = b$, $\mu(\beta) = a$

有向简单弧 Γ^* 的端点按某个顺序对应于参量表示式 (49) 中的参量值 $t = a, b$. 我们把这两个端点区别成 "起点" 和 "终点", 使起点先于终点. 如果 Γ^* 有起点 A 和终点 B, 我们就记为

$$\Gamma^* = \widehat{AB}.$$

于是相反方向的弧就是

$$-\Gamma^* = \widehat{BA}.$$

如果 Γ^* 关于 t 是正向的, 则起点有参量值 a, 而终点有参量值 b.

在一个有向简单弧 $\Gamma^* = \widehat{AB}$ 上按定向一个接着一个地取一系列点 $P_1, \cdots, P_{n-1}$ 可以把它分成有向简单弧段 $\Gamma_1^*, \Gamma_2^*, \cdots, \Gamma_n^*$. 我们令 $P_0 = A$, $P_n = B$, 并且对 $i = 1, \cdots, n$, 定义弧段 Γ_i^*: 它是 Γ^* 上的点集合, 由点 P_{i-1}, P_i 和所有先于 P_i 而跟随 P_{i-1} 的点按 Γ^* 上同样的顺序组成. 我们用符号记作

$$\Gamma^* = \Gamma_1^* + \Gamma_2^* + \cdots + \Gamma_n^*. \tag{51}$$

如果 Γ^* 关于表达式 (49) 中的参量 t 是正向的, 又如果 t_i 是 P_i 所对应的参量值, 则有

$$a = t_0 < t_1 < t_2 < \cdots < t_n = b.$$

当我们限制 t 在区间 $t_{i-1} \leq t \leq t_i$ 上时, 就得到弧段 Γ_i^* (图 1.18).

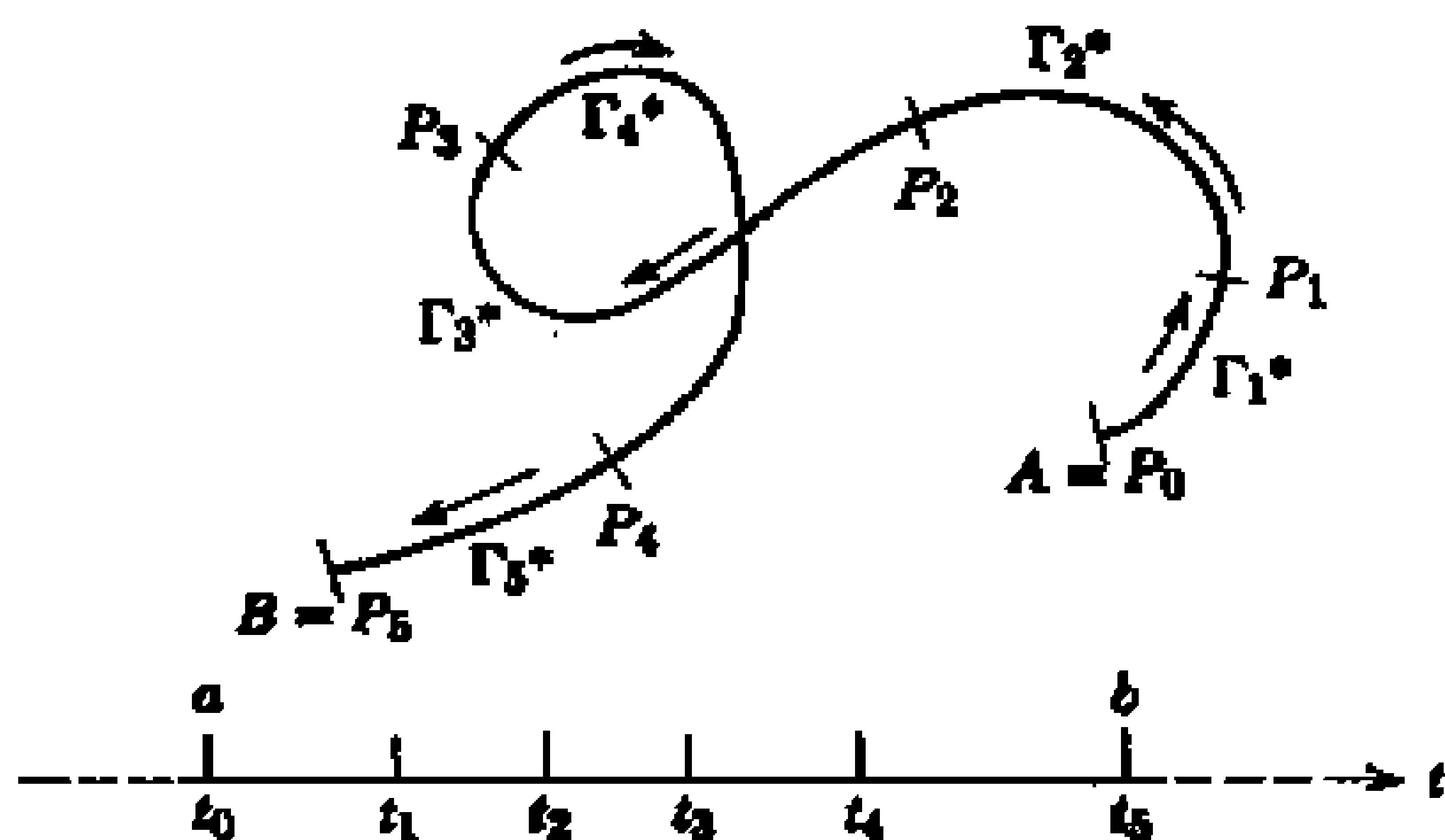


图 1.18 有向弧 $\Gamma^* = AB$ 表示为弧 $\Gamma_{i+1}^* = P_iP_{i+1}$ 的和使 $\Gamma^* = \Gamma_1^* + \Gamma_2^* + \Gamma_3^* + \Gamma_4^* + \Gamma_5^*$.

现在我们可以来定义线性微分型

$$L = A(x,y,z)dx + B(x,y,z)dy + C(x,y,z)dz \tag{52}$$

在一个有向简单弧 Γ^* 上的积分 $\int L$ 了. 我们假设 L 的系数 A, B, C 在 Γ^* 的一个邻域内连续. 我们进一步假设弧 Γ^* 不仅连续并且是**分段光滑**的, 这就是说, 弧 Γ^* 可以用分段光滑的函数 [1]

[1] 这意思是 φ, ψ, χ 在 $a \leq t \leq b$ 上连续, 除去可能有有限个跳跃间断点外, 在这区间内有连续的一阶微商. 注意我们只要求对 Γ^* 存在**某个**分段光滑参量表示式, 而其他表示式无需光滑.

$$x = \psi(t), \quad y = \psi(t), \quad z = \chi(t); \quad a \le t \le b, \tag{53}$$

作为参量表示式.

令 $P_0, P_1, \cdots, P_n$ 是 Γ^* 上任意 $n+1$ 个点，按 Γ^* 所规定的方向排列着，这里 P_0 是 Γ^* 的起点， P_n 是 Γ^* 的终点.

作**黎曼和**

$$F_n = \sum_{\nu=0}^{n-1} (A_\nu \Delta x_\nu + B_\nu \Delta y_\nu + C_\nu \Delta z_\nu), \tag{54}$$

其中 A_ν, B_ν, C_ν 是 A, B, C 在弧 Γ^* 上某一个先于 $P_{\nu+1}$ 而跟随 P_ν 的点 Q_ν 上的值，而 $\Delta x_\nu, \Delta y_\nu, \Delta z_\nu$ 代表

$$x(P_{\nu+1}) - x(P_\nu), \quad y(P_{\nu+1}) - y(P_\nu), \quad z(P_{\nu+1}) - z(P_\nu).$$

我们来证明：当 $n \to \infty$ 时，只要相继的两点 $P_\nu, P_{\nu+1}$ 间的最大距离趋于 0, 这一串 F_n 就收敛到一个极限 F. 这 F 的值不依赖于这些分点 P_ν 或这些中间点 Q_ν 的选取. 我们称 F 为 L 在有向弧 Γ^* 上的积分，并且记作

$$F = \int_{\Gamma^*} L = \int_{\Gamma^*} A dx + B dy + C dz. \tag{55}$$

因为积分的定义中并没有提到参量表示式，显然这个积分不依赖于参量的选取. 存在性的证明就蕴涵着这积分可表为通常的黎曼积分

$$\int_{\Gamma^*} L = \varepsilon \int_a^b \left(A\frac{dx}{dt} + B\frac{dy}{dt} + C\frac{dz}{dt} \right) dt. \tag{56}$$

这里被积分的是一个单变量 t 的函数，它是把 A, B, C 中的变量 x, y, z 用表达式 (53) 代入而得到的；又当 Γ^* 关于 t 的定向是正时，取 $\varepsilon = +1$; 当定向是负的时，取 $\varepsilon = -1$. 同样地，我们也可以把 (56) 写成

$$\int_{\Gamma^*} L = \int_{t_i}^{t_f} \left(A\frac{dx}{dt} + B\frac{dy}{dt} + C\frac{dz}{dt} \right) dt, \tag{57}$$

其中 t_i 是有向弧 Γ^* 的起点的参量值, 而 t_f 是终点的参量值; 这就是说, 当 $\varepsilon=+1$ 时 $t_i=a$, $t_f=b$, 而当 $\varepsilon=-1$ 时 $t_i=b$, $t_f=a$.

为了证明黎曼和 F_n 收敛, 我们利用 Γ^* 的分段光滑的参量表示式 (53). 令 t_ν 是对应于点 P_ν 的参量值. 因为对于简单弧而言, 参量值和曲线上的点之间的对应是双方连续的 (见第 94 页注脚), 所以当 $n\to\infty$ 时相继两点之间的最大距离就趋于 0, 因而 $|t_{\nu+1}-t_\nu|$ 的最大值趋于零. 函数 $\varphi'(t),\psi'(t),\chi'(t)$ 可能在有限个点上有跳跃间断. 我们可以假设所有这些间断点都出现在我们的分点 $t_0,t_1,\cdots,t_n$ 之中, 因为 A,B,C 有界, 并且当 $n\to\infty$ 时, $\Delta x_\nu,\Delta y_\nu,\Delta z_\nu$ 的最大者趋于 0, 所以从黎曼和 F_n 的有限个分点中增加或减少一些是不影响极限的.

现在因为 $\varphi(t),\psi(t),\chi(t)$ 在每个子区间内是可微的, 所以我们可以用微分学中值定理 (见第一卷第 194 页) 得到

$$\Delta x_\nu=\varphi(t_{\nu+1})-\psi(t_\nu)=\varphi'(\tau_\nu)(t_{\nu+1}-t_\nu),$$

$$\Delta y_\nu=\psi'(\tau'_\nu)(t_{\nu+1}-t_\nu),\quad \Delta z_\nu=\chi'(\tau''_\nu)(t_{\nu+1}-t_\nu),$$

其中 $\tau_\nu,\tau'_\nu,\tau''_\nu$ 在 t_ν 和 $t_{\nu+1}$ 之间. Γ^* 上的点 Q_ν 对应的参量值 σ_ν 也在 t_ν 和 $t_{\nu+1}$ 之间. 于是黎曼和 (54) 具有这样的形式

$$F_n=\sum_{\nu=0}^{n-1}[A(\sigma_\nu)\varphi'(\tau_\nu)+B(\sigma_\nu)\psi'(\tau'_\nu)+C(\sigma_\nu)\chi'(\tau''_\nu)]\cdot[t_{\nu+1}-t_\nu].$$

这里点 $t_0,t_1,\cdots,t_n$ 是参量区间 $[a,b]$ 的一个分割. 如果 Γ^* 关于 t 是正向的, 则 t_ν 是一个增序列, $t_0=a$, $t_n=b$, 并且

$$\Delta t_\nu=t_{\nu+1}-t_\nu>0.$$

反之, t_ν 是减序列, $t_0=b$, $t_n=a$, 并且 $\Delta t_\nu<0$. 在我们的参量区间的记号中, a 总代表 a,b 之中较小者, 因而它可以对应于弧 Γ^* 的起点或终点.

定积分是黎曼和的极限，如果我们应用定积分的基本存在定理(见第一卷第 216 页以下)，我们就看到极限 $F=\lim\limits_{n\to\infty}F_n$ 存在并且由公式 (56) 给出 [1]. 因为在定理中假设了用来作黎曼和的所有分点 t_ν 构成一个增序列，所以出现了因子 $\varepsilon=\pm1$. 当 Γ^* 的方向对应于 t 减少时，我们按反方向跑过 t_ν 的值，从 t_n 开始到 t_0 终止，并且改变 Δt_ν 的符号.

显然线积分的定义和公式 (56) 可以推广到 Γ^* 是一个**有向简单闭曲线** [2] 的情况. 在这时为了作 F_n 的表示式 (54) 中的黎曼和我们在 Γ^* 上按定向规定的次序一个接着一个选 n 个点 $P_1,P_2,\cdots,P_n$，并且 $P_0=P_n$.

(x,y) 平面上的曲线积分的例子在第一卷中已经遇到过. 例如把一有向闭曲线 Γ^* 所围的有向面积表示成

$$A=\frac{1}{2}\int_a^b\left(x\frac{dy}{dt}-y\frac{dx}{dt}\right)dt$$

(见第一卷第 411 页); 它写成线积分是

$$A=\frac{1}{2}\int_{\Gamma^*}xdy-ydx.$$

另一个例子是力场作功 W，力场的分量是 ρ,σ，运动沿着曲线 $\Gamma^*=\widehat{P_0P_1}$ 从点 P_0 到点 P_1，取弧长 s 作参量. 这里 (见第一卷第 470 页)

$$W=\int_{s_0}^{s_1}\left(\rho\frac{dx}{ds}+\sigma\frac{dy}{ds}\right)ds,$$

它可以写成

$$W=\int_{\Gamma^*}\rho dx+\sigma dy.$$

---

1) 为了收敛性不需要中间值 $\tau_\nu,\tau'_\nu,\tau''_\nu,\sigma_\nu$ 相同 (见第一卷第 219 页的附注).

2) 这种曲线有一个连续参量表示式 (53)，除去 $t=a$ 和 $t=b$ 对应到同一个点外，不同的 t 对应于不同的点. 此外，在 Γ^* 上有一个指定的循环顺序，它对应于 t 增加或减少 (见第一卷第 386 页). 我们总可以把 Γ^* 表示成有向简单弧 Γ_i^* 的和，如像 (51) 式的形式，对于 $i=2,\cdots,n$，Γ_{i-1}^* 的终点是 Γ_i^* 的起点，而 Γ_n^* 的终点是 Γ_1^* 的起点.

我们可以用同样的方法把分量为 ρ, σ, τ 的空间力场，沿一个弧 Γ^* 按有向弧自己定向的方向移动所作的 功 定义为一个曲线积分：

$$W = \int_{\Gamma^*} \rho dx + \sigma dy + \tau dz.$$

练 习 1.9 b

1. 求

$$\int zdx + xdy + ydz$$

(a) 沿螺旋线

$$x = \cos t, \quad y = \sin t, \quad z = t$$

联接点 $(1,0,0)$ 和点 $(1,0,2\pi)$ 的一段弧；

(b) 沿抛物线

$$x = x_0(1-t^2), \quad y = y_0(1-t^2), \quad z = t$$

联接点 $(0,0,1)$ 和 $(0,0,-1)$ 的一段弧 (x_0, y_0 为常数).

c. 线积分对端点的相关性

我们回到由 (52) 给出的一般的微分型 L. 设 Γ 是一个简单弧 (尚未定向), 有分段光滑的参量表示式 (53).

对于 Γ 上任意两点 P_0, P_1, 取它们对应的参量 t 的值 t_0, t_1, 我们可以按公式 (57) 作积分

$$I = \int_{t_0}^{t_1} \left(A\frac{dx}{dt} + B\frac{dy}{dt} + C\frac{dz}{dt}\right)dt.$$

I 等于沿 Γ 以 P_0 为起点， P_1 为终点的一段有向弧 $\overset{\frown}{P_0P_1}$ 的积分 $\int L$. 因此 I 与参量表示式无关. 我们记作

$$I = \int_{P_0}^{P_1} L.$$

I 的值是由一对有序的点 P_0, P_1 和以它们为端点的简单弧所确定的.

对于固定的 P_0, 沿着 弧 Γ 我们可以用下面的不定积分来定义一个函数 $f = f(P)$:

$$f(P) = \int_{P_0}^{P} L = \int_{t_0}^{t} \left(A\frac{dx}{dt} + B\frac{dy}{dt} + C\frac{dz}{dt}\right)dt. \tag{58}$$

把 f 作为自变量 t 的函数, 我们有

$$\frac{df}{dt} = A\frac{dx}{dt} + B\frac{dy}{dt} + C\frac{dz}{dt}. \tag{59}$$

把这个等式写成

$$df = \frac{df}{dt}dt = Adx + Bdy + Cdz = L,$$

于是就把线性微分型表示成一个函数 f 的微分了 (不一定是恰当微分); 但是我们要记住这个关系式只是沿着定义 f 用的一条特殊的曲线 Γ 才成立.

如果我们把线积分表示为对变量 t 的积分并且应用定积分和不定积分的基本关系 (见第一卷第 213 页), 就立即得到: 对 Γ 上的任意两点 P 和 P' 有

$$\int_{P}^{P'} L = f(P') - f(P). \tag{60}$$

特别地, 如果 Γ 确定了某个方向后成为 Γ^*, 它以 A 为起点 B 为终点, 则有

$$\int_{\Gamma^*} L = \int_{A}^{B} L = f(B) - f(A). \tag{61}$$

若 $P_0, \cdots, P_n$ 是 Γ^* 上的点, 它们的顺序按 Γ^* 的定向, $P_0 = A$, $P_n = B$, 那么我们有

$$\int_{\Gamma^*} L = f(B) - f(A) = \sum_{\nu=0}^{n-1}[f(P_{\nu+1}) - f(P_\nu)] = \sum_{\nu=0}^{n-1}\int_{P_\nu}^{P_{\nu+1}} L.$$

如果我们用 $\Gamma^*_{\nu+1}$ 表示以 P_ν 为起点 $P_{\nu+1}$ 为终点的一段弧，就有

$$\int_{P_\nu}^{P_{\nu+1}} L = \int_{\Gamma^*_{\nu+1}} L.$$

这里 Γ^*_ν 的方向与 Γ^* 的一致，从而

$$\Gamma^* = \Gamma^*_1 + \Gamma^*_2 + \cdots + \Gamma^*_n.$$

所以**线积分是可加的**:

$$\int_{\Gamma^*_1+\cdots+\Gamma^*_n} L = \int_{\Gamma^*_1} L + \cdots + \int_{\Gamma^*_n} L. \tag{62}$$

类似地，如果我们交换 Γ^* 的端点，便有

$$\int_{-\Gamma^*} L = -\int_{\Gamma^*} L. \tag{63}$$

把这些法则运用到有向闭曲线被表示成为有向简单弧的和的时候，特别有用．试考虑几个可以有公共部分的有向简单闭曲线 $C^*_1, \cdots, C^*_n$ (见图 1.19).

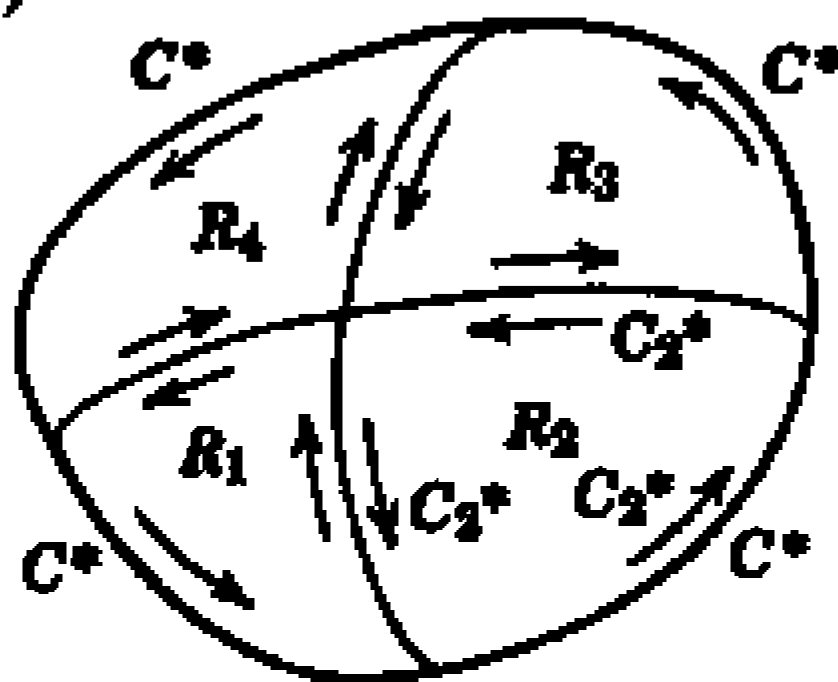


图 1.19　闭曲线上的线积分的可加性

设简单弧 Γ 是曲线 C^*_i 和 C^*_k 所共有的，它由 C^*_i 和 C^*_k 所得到的方向相反，并设曲线 $C^*_1, \cdots, C^*_n$ 上不被任何两曲线共有的部分加起来是一个有向闭曲线 C^*. 把曲线 C^*_i 上的线积分写成简单弧上的积分之和，把所有这些简单弧上的积分都加起来，那么每段公共弧上的两个积分恰好相消，我们就得到公式

$$\int_{C^*} L = \int_{C^*_1} L + \cdots + \int_{C^*_n} L. \tag{64}$$

特别地，当 C_i^* 是平面曲线，是二维的互不重迭的区域 R_i 的边界，所有 R_i 的和是区域 R, R 的边界曲线是 C^*, 所有 C_i^* 和 C^* 有相同的方向时，就出现上面的情况. 更一般地，区域 R 和它的边界 C^* 可以是在一张曲面上，并且 R 被一些弧分割成一些子区域 R_i, R_i 的边界曲线是 C_i^*, 所以 C_i^* 的方向恰为所描述的式样时，也有上面的结果.

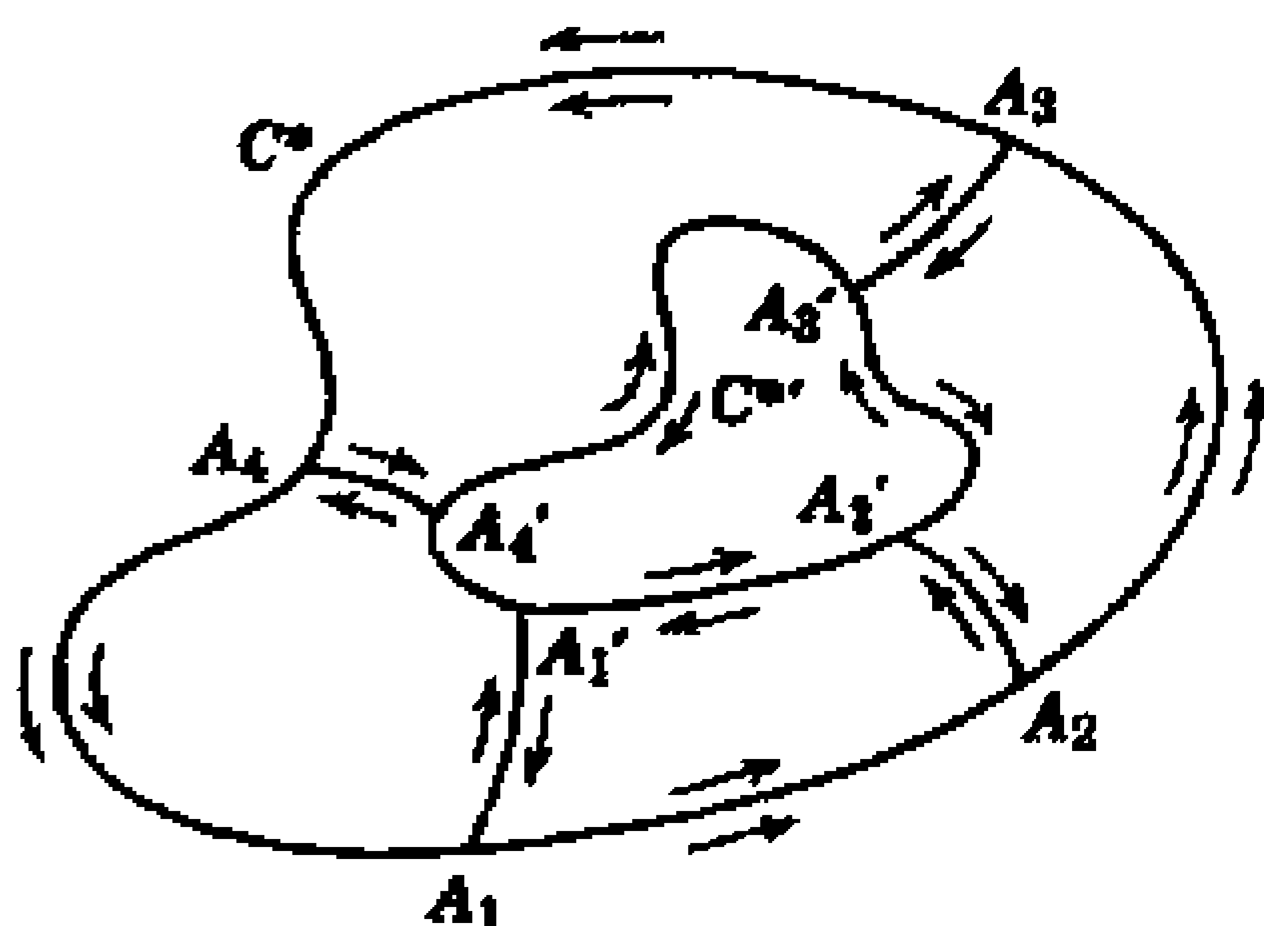


图 1.20

下面的定理是这原理的另一应用. 设 C^* 和 C'^* 是两个有向闭曲线 (见图 1.20), 分别按定向的方向被两组点 $A_1, \cdots, A_n$ 和 $A_1', \cdots,$ 分割，并设每一对对应点 A_i 和 A_i' 被一曲线连接. 如果用 C_i^* 记有向闭曲线 $A_iA_{i+1}A_{i+1}'A_i'$ (把 A_{n+1} 等同于 A_1, A_{n+1}' 等同于 A_1'), 则

$$\sum_{i=1}^{n} \int_{C_i^*} L = \int_{C^*} L - \int_{C'^*} L. \tag{65}$$

1.10 线性微分型的可积性的基本定理

a. 全微分的积分

线性微分型

$$L = Adx + Bdy + Cdz \tag{66}$$

中特别重要的一类是函数 $u=f(x,y,z)$ 的全微分，其中 A,B,C 有如下的形式

$$A=\frac{\partial f}{\partial x},\quad B=\frac{\partial f}{\partial y},\quad C=\frac{\partial f}{\partial z},\tag{67}$$

函数 f 有连续的一阶微商. 虽然在一般情况下，$\int_{\Gamma^*}L$ 的值不仅与端点有关并且与曲线的整个路径有关，但在这里，下面的定理是成立的：

当线性微分型 L 是一个函数 f 的全微分时，它的积分等于 f 在两个端点上的值的差，而与端点间的曲线 Γ^* 无关. 这就是说，对一切曲线 Γ^*，只要它完全位于 f 的定义域里并且有同一个起点 P_0 和同一个终点 P_1，那么 $\int_{\Gamma^*}L$ 的值就相同.

为了证明，令曲线 Γ^* 参照一个参量 t, t_0 对应于起点 P_0, t_1 对应于终点 P_1. 由第 97 页公式 (57) 有

$$\int_{\Gamma^*}L=\int_{t_0}^{t_1}\left(A\frac{dx}{dt}+B\frac{dy}{dt}+C\frac{dz}{dt}\right)dt.$$

根据微商的锁链法则 (见第 57 页公式 (18)) 我们有

$$\int_{\Gamma^*}L=\int_{t_0}^{t_1}\frac{df}{dt}dt=f\Big|_{t_0}^{t_1}=f(P_1)-f(P_0),\tag{68}$$

其中

$$f(P_i)=f(x(t_i),y(t_i),z(t_i)),\quad i=0,1.$$

我们注意，积分与路径无关的要求可以由在简单闭曲线 Γ^* 上积分值为 0 的要求来代替，这是因为如果我们用两个点 P_0 和 P_1 把曲线 Γ^* 分割成两个有向弧 Γ_1^* 和 Γ_2^*，Γ_1^* 以 P_0 为起点 P_1 为终点，Γ_2^* 以 P_1 为起点 P_0 为终点，就有 (见第 96 页)

$$\Gamma^*=\Gamma_1^*+\Gamma_2^*.$$

于是

$$\int_{\Gamma^*} L = \int_{\Gamma_1^*} L + \int_{\Gamma_2^*} L = \int_{\Gamma_1^*} L - \int_{-\Gamma_2^*} L.$$

这里 $-\Gamma_2^*$ 与 Γ_1^* 有相同的起点 P_0 和相同的终点 P_1. 所以在闭曲线 Γ^* 上 $\int L$ 为零的确与 L 在两个有相同起点 P_0 和相同终点 P_1 的简单弧上积分相等是一样的.

b. 线积分只依赖于端点的必要条件

只是在非常特殊的条件下线积分才与路径无关, 也就是说, 沿闭路的线积分才为零. 例如, 若 (x,y) 平面上一条闭曲线 C^* 作为边界所包围的区域有正的面积, 那么在 C^* 上线积分 $\int xdy - ydx$ 就不是零. 在前一节中我们证明了: $\int L$ 与联接端点的路径无关的一个充分条件是 L 为一个全微分. 现在线积分理论的主要任务便是来证明这个条件也是必要的, 然后把这个充要条件表示成一个便于应用的形式.

我们将对三维空间曲线上的积分来研究这个无关性问题. 但是结论和证明对任何维数都是完全类似的. 我们假设

$$L = Adx + Bdy + Cdz$$

是一个线性微分型, 系数 A, B, C 是 x, y, z 在空间的一个开集 R 内的连续函数. 则下面的定理成立:

线积分 $\int L$ 在 R 内的一条有向简单弧 Γ^* 上的值与 Γ^* 的特殊选择无关, 而只决定于 Γ^* 的起点和终点, 当且仅当 L 是 R 内的一个函数 $f(x,y,z)$ 的全微分.

在第 104 页上我们已经证明了这个条件是充分的; 也就是说, 对于一个恰当微分 $L = Adx + Bdy + Cdz$, 线积分 $\int L$ 与路径无关. 不难看出这个条件也是必要的. 设 $\int_{\Gamma^*} L$ 只依赖于 Γ^* 的端点. 我们要证明的是: 存在一个定义在 R 内的函数 $u(x,y,z)$ 使得

$du = L$. 不失一般性，我们可以假设 R 内任意两个点可以用完全在 R 内的简单折线连接起来 [1]. 我们在 R 内取一个固定点 P_0, 定义 R 上任一点 P 处函数 $u = u(x, y, z) = u(P)$ 为以 P_0 为起点、 P 为终点的任一简单弧上的积分 $\int L$. 为了计算 u 的偏微商，考虑 R 的任意一点 $(x, y, z) = P$ (图 1.21). 因为 R 是开的，只要 $|h|$ 充分

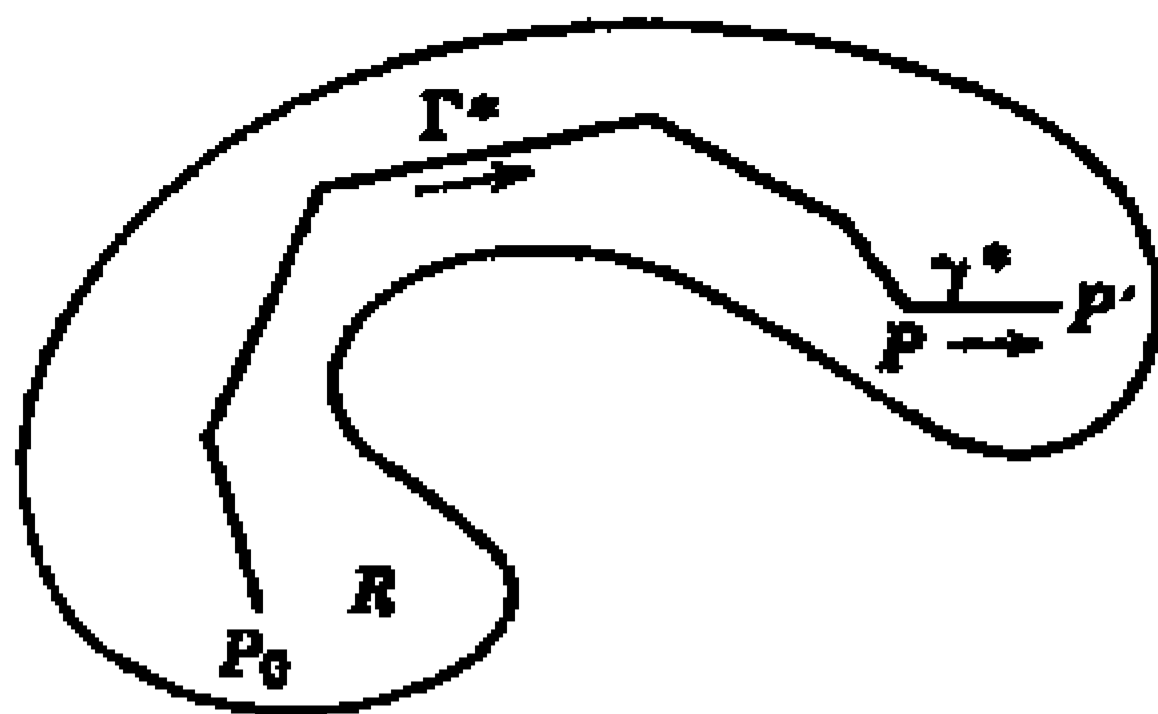


图 1.21

小，所有点 $(x + h, y, z) = P'$ 就也属于 R. 令 γ^* 是连接 P 到 P' 的有向直线段， Γ^* 是从 P_0 到 P 的简单折线. 我们总可以稍微修改 Γ^* 使得以 P 为终点的折线的最后一段不与 x 轴平行. 那么 Γ^* 与 γ^* 除去 P 之外没有公共点 (至少对 $|h|$ 充分小是这样), 而 $\Gamma^* + \gamma^*$ 表示一个以 P_0 为起点、 P' 为终点的简单弧. 于是 (见第 102 页 (62)) 有

$$\begin{aligned} &u(x+h, y, z) - u(x, y, z) \\ &\qquad = u(P') - u(P) = \int_{\Gamma^*+\gamma^*} L - \int_{\Gamma^*} L \\ &\qquad = \int_{\gamma^*} L = \int_x^{x+h} A(t, y, z)dt. \end{aligned}$$

用 h 除并令 $h \to 0$ 取极限，我们的确得到

$$\frac{\partial u(x, y, z)}{\partial x} = A,$$

---

1) 开集 R 总可以分解成一些连通的子集各具有这个性质 (见附录 120). 然后我们可在每一个这样的子集上用下面指出的构造方法定义 u.

还可以类似地得到 $\frac{\partial u}{\partial y}=B$, 和 $\frac{\partial u}{\partial z}=C$. 这就证明了 $du=L$.

c. 可积条件的不足

除非我们有办法判别一个给定的微分型 L 是不是全微分, 否则我们刚才证明的线积分与路径无关的定理是没有多大价值的. 我们需要一个只涉及 $L=Adx+Bdy+Cdz$ 的系数 A,B,C, 并且便于检验的条件. 我们已经知道可积条件

$$\frac{\partial B}{\partial z}-\frac{\partial C}{\partial y}=0,\quad \frac{\partial C}{\partial x}-\frac{\partial A}{\partial z}=0,\quad \frac{\partial A}{\partial y}-\frac{\partial B}{\partial x}=0 \tag{69}$$

是存在一个函数 $u=f(x,y,z)$ 使 $L=du$ 的必要条件. 我们把满足条件 (69) 的微分型 L 称为**闭的**. 那么每一个恰当微分型都是闭的. 因为只有当 L 是一个全微分时线积分才能够与联结两点的路径无关, 所以我们看到, **如果要** $\int L$ **只依赖于积分路径的端点, 那么条件**(69)**是必要的**. 现在的问题是: 这些条件是否也是充分的呢? 如果这些条件使我们可以构造一个函数 $u=f(x,y,z)$ 使得

$$A=\frac{\partial f}{\partial x},\quad B=\frac{\partial f}{\partial y},\quad C=\frac{\partial f}{\partial z}, \tag{70}$$

那么这些条件就是充分的. 意外的是, 对于保证 L 是某一个函数 u 的全微分, 从而保证 $\int L$ 与路径无关来说, 可积条件 (69) 几乎是足够的, 但又是不完全的. 仅仅有等式 (69) 是不够的, 但是如果我们对 L 所在的空间区域加上相当特殊的**几何性质**方面的假设, 那就够了.

一个简单的反例指出, 单有条件 (69) 不足以保证 $\int L$ 在任一闭曲线上为零. 试考虑微分型

$$L=\frac{xdy-ydx}{x^2+y^2}, \tag{71}$$

这相当于选择系数

$$A=\frac{-y}{x^2+y^2},\quad B=\frac{x}{x^2+y^2},\quad C=0,$$

除了直线 $x=y=0$ (即 z 轴) 上的点之外，它们处处都有定义. 容易验算可积条件 (69) 是满足的，因而 L 是闭的. 当我们沿着 (x,y) 平面上的单位圆 $C^*: x=\cos t,\ y=\sin t,\ z=0$, 按照关于 t 的正方向积分时，就得到

$$\begin{aligned}\int_{C^*} L &= \int_0^{2\pi}\left(A\frac{dx}{dt}+B\frac{dy}{dt}\right)dt = \int_0^{2\pi}(\sin^2 t+\cos^2 t)dt\\ &= 2\pi \neq 0.\end{aligned}$$

其实，容易计算沿着任一闭曲线 C 的 $\int L$. 引进点 $P(x,y,z)$ 的极角 θ:

$$\cos\theta=\frac{x}{\sqrt{x^2+y^2}},\quad \sin\theta=\frac{y}{\sqrt{x^2+y^2}}. \tag{72}$$

这个角 θ 是由 (x,z) 平面和过点 P 与 z 轴的平面所构成的 (见图 1.22). 于是

$$d\theta = d\arctan\frac{y}{x}=L, \tag{73}$$

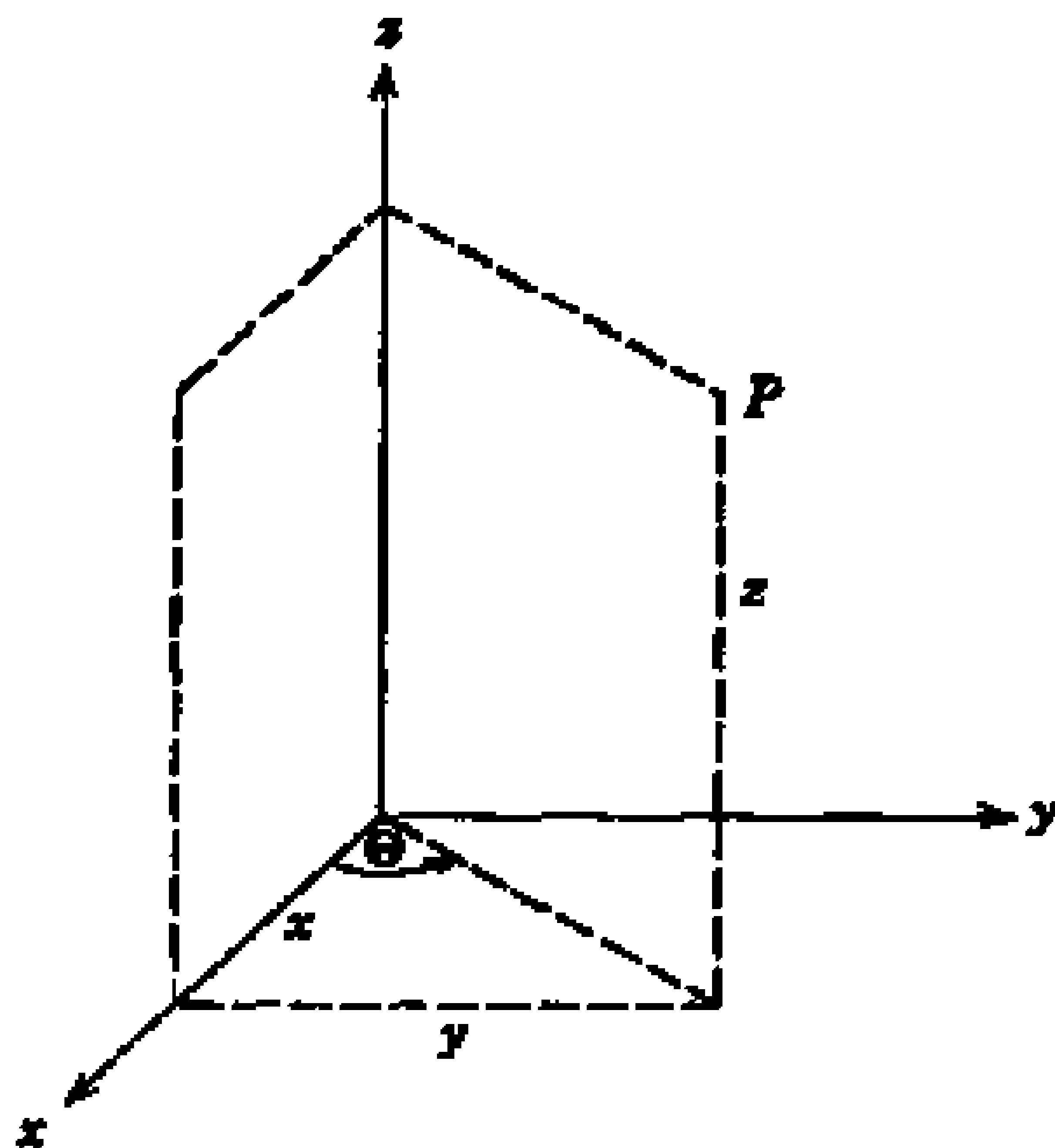


图 1.22

所以 L 被表成了 $u=\theta$ 这函数的全微分. 问题复杂在于由公式 (72) 所确定的 θ 的值可以差 2π 的倍数. 在一点 P_0 处从 θ 的某个可能

的值 θ_0 出发，我们可以定义 θ 在任一点 P 的值，办法是用一连续曲线连接点 P_0 到 P，并令

$$\theta(P)=\theta_0+\int_{P_0}^{P}d\theta=\theta_0+\int L$$

(见第一卷第 484 页)．但是这样定义的 $\theta(P)$ 是多值的，依赖于曲线的选取：因为对于一个闭曲线 C^*，

$$\frac{1}{2\pi}\int_{C^*}d\theta$$

表示曲线 C^* 按反时针方向绕 z 轴的次数 (见图 1.23)．

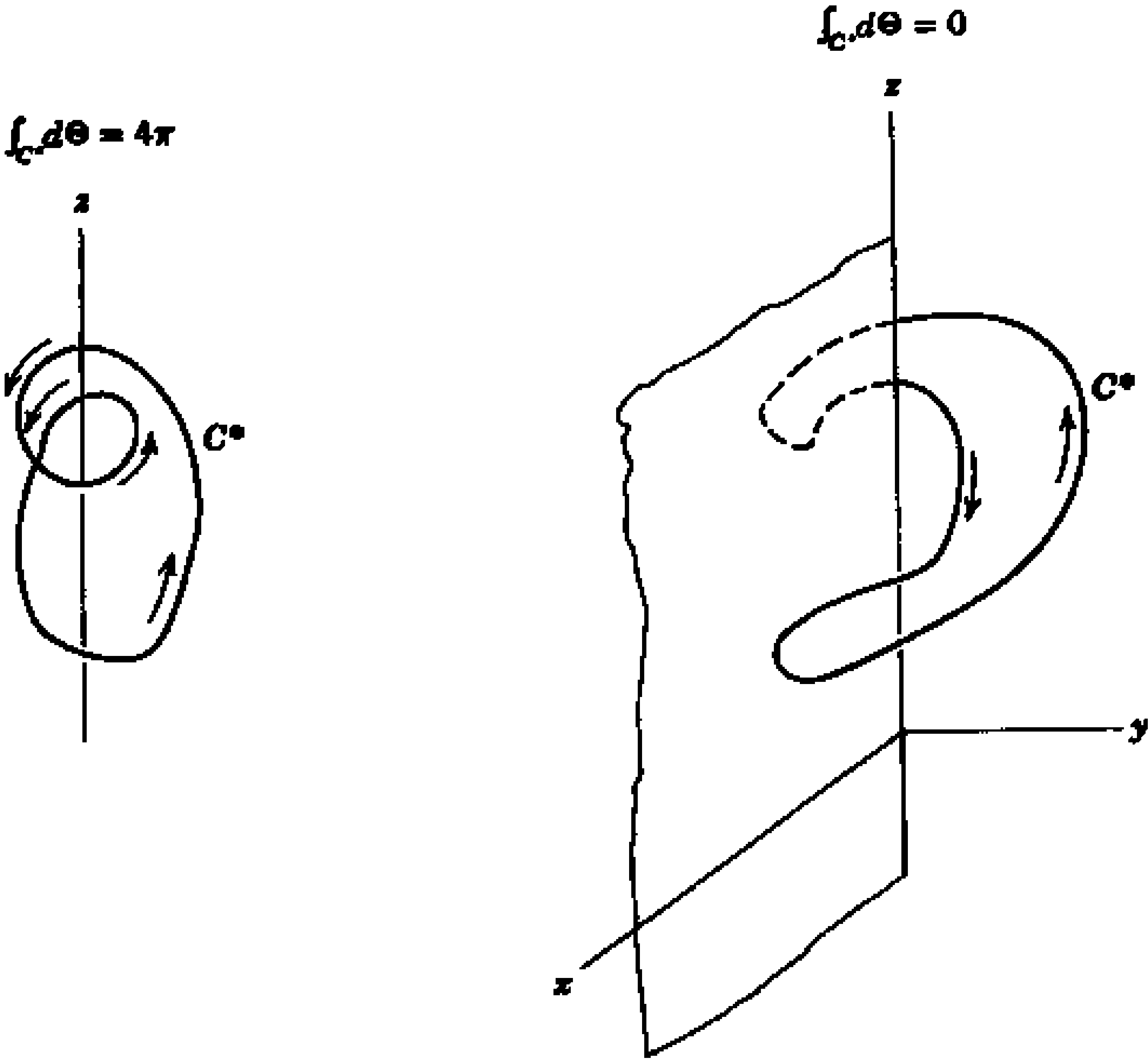


图 1.23

所以，沿着以 P_0, P 为端点的两条不同的路径，若要积分

$$\int_{P_0}^{P}d\theta \tag{74}$$

的值相等，就只有当从 P_0 沿一条路径到 P 再沿另一条路径回到 P_0 时，闭曲线绕 z 轴零次．而我们是可以阻止任何路径围绕 z 轴的；这只需要求点 $P(x,y,z)$ 有 $y\neq 0$，或者当 $y=0$ 时 $x>0$；或者这样来说，沿着半平面

$$y=0,\quad x\leq 0$$

竖起一座墙，不许通过．没有除去的点形成一个区域 R，在其中我们可以指定 θ 以唯一的值，并有

$$-\pi<\theta<\pi.$$

这样就构造了一个连续可微函数 $\theta=\theta(x,y,z)$，它的微分是 L．在这个区域里，沿联接 P_0 到 P 的任一路径，积分 (74) 都有唯一的值 $\theta(P)-\theta(P_0)$，与路径无关．同样地，沿这个区域里的任何一条闭路，积分值都为 0．

d. 单连通集

为了一般地确切叙述基本定理，我们需要**单连通**[1]**开集**的概念．一个单连通集 R 是这样一个集合：R 内任意两点可以用一条完全位于 R 内的路径连接起来，并且 R 内任意两条有相同端点的路径可以不移动端点而始终在 R 内互相变形．

我们给出上述概念的确切的定义如下．连接两点 $P'=(x',y',z')$ 和 $P''=(x'',y'',z'')$ 的一条路径 C 在 R 内的意思是：有三个连续函数 $\varphi(t),\psi(t),\chi(t)$ 定义在区间 $0\leq t\leq 1$ 上，使得对于该区间内一切 t，点 $P(t)=(\varphi(t),\psi(t),\chi(t))$ 都位于 R 内，并且 $t=0$ 时与 P' 重合，$t=1$ 时与 P'' 重合[2]．集合 R 叫做**连通的**[3]如果 R 内

1) 更确切地说是“按路径单连通”．

2) 不需要不同的 t 对应到不同的 $P(t)$．注意一条路径的描述不只包括空间中点 $P(t)$ 的集合 (路径的“支集”)，也还包括相应的参量 t 的选取．空间中每一个简单弧确定许多不同的路径，它们对应于这个弧的不同的参量表示式．我们总可以经过线性变换让参量值在区间 $0\leq t\leq 1$ 上变化．

3) 更确切地说是“按路径连通”．

任意两点 P' 和 P'' 都可以用 R 内的一条路径连接起来. 事实上容易看出, 只要集合 R 是开的, 它们也就可以用 R 内光滑的简单弧连接起来[1].

连通集的最简单的例子是**凸集** R. 凸集 R 的特征性质是: 其中任意两点 P' 和 P'' 可以用 R 内的一个直线段连接. 这样我们可以给端点是 $P'=(x',y',z')$ 和 $P''=(x'',y'',z'')$ 的线性路径简单地选取三个线性函数

$$\varphi(t)=(1-t)x'+tx'',\quad \psi(t)=(1-t)y'+ty'',$$

$$\chi(t)=(1-t)z'+tz'',$$

其中 $0\leq t\leq 1$. 这种凸集的例子有球体、立方体. 连通但并不凸的集合的例子有圆环体、球壳 (也就是两个同心球之间的空间) 和球或柱的外部. 空间内任何一个不连通的集合 R 都是由一些连通子集组成的, 这些子集称为 R 的**分支**. 不连通集的例子有: 不属于一个球壳的点所成的集合, 没有整数坐标的点所成的集合.

令 C_0 和 C_1 是 R 内任意两条路径, 分别由 $(\varphi_0(t),\psi_0(t),\chi_0(t))$ 和 $(\varphi_1(t),\psi_1(t),\chi_1(t))$ 给出, 它们有共同的端点 P',P'', 对应于 $t=0$, $t=1$. 如果我们可以用以 P',P'' 为共同端点的连续的一族路径 C_λ "把 C_0 变形为 C_1" 或 "联接 C_0 和 C_1", 那么这个连通集 R 就是单连通的. 这个意思是, 存在着两个变量 t,λ 在 $0\leq t\leq 1, 0\leq\lambda\leq 1$ 上的连续函数 $(\varphi(t,\lambda),\psi(t,\lambda),\chi(t,\lambda))$, 使得点 $P=(\varphi,\psi,\chi)$ 总在 R 内, 并且 $\lambda=0$ 时 P 与 $(\varphi_0,\psi_0,\chi_0)$ 重合, $\lambda=1$ 时与 $(\varphi_1,\psi_1,\chi_1)$ 重合, $t=0$ 时与 P' 重合, $t=1$ 时与 P'' 重合[2]. 对于每一固定的 λ, 函数 φ,ψ,χ 给出了 R 内联结点 P' 和 P'' 的一条路径 C_λ. 当 λ 从 0 变到 1 时, 路径 C_λ 从 C_0 连续地变到 C_1, 这样就说是有一个从 C_0 到 C_1 的 "连续变形"(见图 1.24).

---

1) 取参量区间的一个足够细的分割, 用直线段联接相应的点 $P(t)$, 就得到 R 内的一个联接 P' 到 P'' 的折线弧. 取消掉圈子, 就得到一个简单折线弧. 在靠近尖角的一小部分, 用适当的抛物线弧替换, 就得到 R 内的一条联接 P' 到 P'' 的光滑简单弧. 再参看第 121 页.

2) 路径 C_0 和 C_1 叫做关于 P',P'' **同伦**.

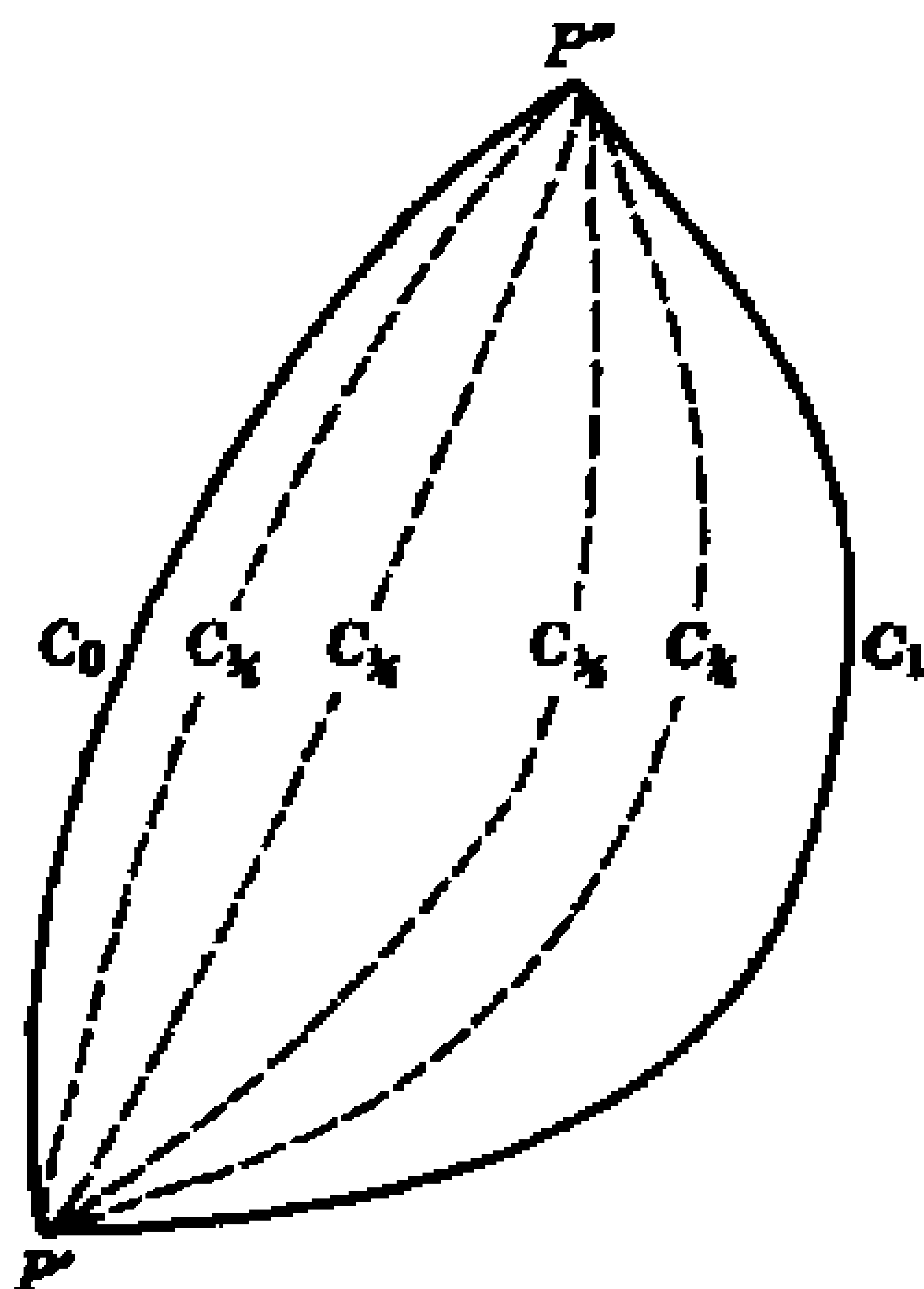


图 1.24

容易看出，凸集 R 是单连通的. 这只要我们把以 P', P'' 为共同端点的两条路径 C_0, C_1 用路径 C_λ:

$$\varphi(t,\lambda)=(1-\lambda)\varphi_0(t)+\lambda\varphi_1(t),$$

$$\psi(t,\lambda)=(1-\lambda)\psi_0(t)+\lambda\psi_1(t),$$

$$\chi(t,\lambda)=(1-\lambda)\chi_0(t)+\lambda\chi_1(t),$$

联结起来. 这里的 C_λ 可以用几何方法得到: 把 C_0 和 C_1 上对应于同一个 t 的点用直线段连接起来，并取线段上按比例 $\lambda/(1-\lambda)$ 的分点作为 C_λ 的相应点. 由于 R 的凸性，用这种方法得到的这些点都在 R 内. 球壳是另一种不同类型的按路径单连通的集合. 另一方面，从 (x,y,z) 空间撤去 z 轴所得的集合 R 是不单连通的. 在这里两条路径 (半圆)

$$x=\cos\pi t,\quad y=\sin\pi t,\quad z=0\quad(0\leq t\leq 1)$$

和

$$x=\cos\pi t,\quad y=-\sin\pi t,\quad z=0\quad(0\leq t\leq 1)$$

有相同的端点，但是不能互相变形而不经过不属于 R 的 z 轴[1]。

e. **基本定理**

现在我们可以来叙述**闭的**和**恰当的**微分型之间的关系了。

如果微分型 $L = Adx + Bdy + Cdz$ **的系数在一个单连通集** R **内有连续的一阶偏微商并且满足可积条件**

$$B_z - C_y = 0, \quad C_x - A_z = 0, \quad A_y - B_x = 0, \tag{75a}$$

则 L **是定义在** R **内的一个函数** u **的全微分**：

$$A = u_x, \quad B = u_y, \quad C = u_z. \tag{75b}$$

为了证明这个结论，只要证明 L 沿着 R 内任何一条以 P' 为起点、P'' 为终点的简单折线的积分值都只依赖于 P' 和 P'' (见第 105 页)。我们把 C_0^* 和 C_1^* 这两段有向弧分别表示成参量形式：

$$x = \varphi_0(t), \quad y = \psi_0(t), \quad z = \chi_0(t) \quad (0 \le t \le 1); \tag{76a}$$

$$x = \varphi_1(t), \quad y = \psi_1(t), \quad z = \chi_1(t) \quad (0 \le t \le 1); \tag{76b}$$

$t = 0$ 时得到 P'，$t = 1$ 时得到 P''。应用 R 的单连通性，我们可以在路径 (76a, b) 中"嵌入"连续的路径族[2]

$$x = \varphi(t, \lambda), \quad y = \psi(t, \lambda), \quad z = \chi(t, \lambda); \tag{76c}$$

当 $\lambda = 0, 1$ 时就成为 (76a, b)，而当 $t = 0, 1$ 时就成为 P', P''。由第 90 页公式 (56) 我们有

$$\int_{C_1^*} L - \int_{C_0^*} L = \int_0^1 [(Ax_t + By_t + Cz_t)|_{\lambda=1}$$

---

1) 这可以从下面的基本定理和这样一个事实得到，这事实就是存在着闭微分型 ((71) 就是一个)，它在这个圆周上的积分不是零。

2) 这族中对于 $\lambda \neq 0, 1$ 的路径不需要是简单的。

$$-(Ax_t+By_t+Cz_t)|_{\lambda=0}]dt, \tag{76d}$$

其中 x,y,z 是 (76c) 所给出的 t,λ 的函数. 首先，我们假设这些函数对 t,λ 有连续的一阶偏导数，并且在 $0\le t\le 1, 0\le\lambda\le t$ 上有连续的混合二阶偏导数. 于是由 (76d) 得到

$$\int_{C_1^*}L-\int_{C_0^*}L=\int_0^1 dt\int_0^1(Ax_t+By_t+Cz_t)_\lambda d\lambda. \tag{76e}$$

现在用微商的锁链法则和可积条件 (75a) 得到等式

$$\begin{aligned}(Ax_t+By_t+Cz_t)_\lambda=&Ax_{\lambda t}+By_{\lambda t}+Cz_{\lambda t}+A_xx_\lambda x_t\\&+A_yy_\lambda x_t+A_zz_\lambda x_t+B_xx_\lambda y_t+B_yy_\lambda y_t\\&+B_zz_\lambda y_t+C_xx_\lambda z_t+C_yy_\lambda z_t+C_zz_\lambda z_t\\=&(Ax_\lambda+By_\lambda+Cz_\lambda)_t.\end{aligned}$$

交换积分顺序 (见第 88 页), 我们有

$$\int_{C_1^*}L-\int_{C_0^*}L=\int_0^1 d\lambda\int_0^1(Ax_\lambda+By_\lambda+Cz_\lambda)_t dt;$$

而因为端点不依赖于 λ, 所以 $x_\lambda,y_\lambda,z_\lambda$ 当 $t=0,1$ 时为零，于是得到

$$\int_{C_1^*}L-\int_{C_0^*}L=0.$$

我们看到关于 R 为单连通的假设在证明中所起的重要作用，它使我们能够把线积分化为 R 中某个区域上的一个重积分.

容易取消掉要求函数 φ,ψ,χ 的各个微商存在的限制. 可以只假设弧 C_0^* 和 C_1^* 是光滑的，这就是说，函数 $\varphi(t,\lambda),\psi(t,\lambda),\chi(t,\lambda)$ 当 λ 取值 0 或 1 时对 t 有连续的微商，而 λ 取其他的值时是连续的. 于是我们可以用对 t 和 λ 有连续的一阶偏微商并且有混合二阶偏微商的函数 $\overline{\varphi},\overline{\psi},\overline{\chi}$ 来一致地逼近以上各函数 (见第 90 页). 为了使所得到的更为光滑的函数表示 C_0^* 和 C_1^* 之间的一个变形，它

们当 $\lambda=0,1$ 时和 $t=0,1$ 时都应与 φ,ψ,χ 一致. 这个要求总是可以达到的, 只要稍微修改 $\overline{\varphi},\overline{\psi},\overline{\chi}$, 加上一些适当的项:

$$\begin{aligned}x=&\overline{\varphi}(t,\lambda)-(1-\lambda)[\overline{\varphi}(t,0)-\varphi_0(t)]-\lambda[\overline{\varphi}(t,1)-\varphi_1(t)]\\&-(1-t)[\overline{\varphi}(0,\lambda)-\varphi_0(0)]-t[\overline{\varphi}(1,\lambda)-\varphi_0(1)]\\&+(1-t)(1-\lambda)[\overline{\varphi}(0,0)-\varphi_0(0)]+(1-t)\lambda[\overline{\varphi}(0,1)\\&-\varphi_0(0)]+t(1-\lambda)[\overline{\varphi}(1,0)-\varphi_0(1)]+t\lambda[\overline{\varphi}(1,1)-\varphi_0(1)];\end{aligned}$$

对 y 和 z 也有类似的式子. 这些函数当 $\lambda=0,1$ 时和 $t=0,1$ 时都取得应有的值, 还有连续的一阶偏微商和混合二阶偏微商, 并且可以逼近原来的函数 φ,ψ,χ 如此接近以至相应的点 (x,y,z) 也在开集 R 内.

最后, 这 L 的积分等式可以沿着只是*逐段*光滑的弧 C_0^*,C_1^* (例如折线) 来进行, 只要用有相同的端点的光滑的弧来逼近这些弧就行了. 在光滑的弧上积分值都相同, 从而在 C_0^* 和 C_1^* 上的积分值作为极限也是相同的.

附　　录

几何直观和物理实际经常给创造性的数学思维提供有力的启发和指导. 然而由于 19 世纪初以来数学分析的发展, 停止借助于直观作为数学考虑的主要依据, 已经成为十分迫切的需要. 人们愈来愈致力于严密的证明, 以公理化加强了的确切性、和叙述清楚的概念和推理步骤为基础. 在这个发展过程中*集合*, 特别是*点集*的概念起了很重要的作用, 而现在它们已被吸收到分析的结构里来了. 这个附录将对这个发展中的某些东西作一个简单的介绍.

A.1　多维空间的聚点原理及其应用

我们可以用与一元函数的情况下完全同样的方法把多元函数

理论建立在一个坚实的基础上. 只要对两个变量的情况来讨论就可以了, 因为对自变量多于两个的函数而言, 这些方法基本相同.

a. 聚点原理

我们讨论的基础是波尔扎诺和魏尔斯特拉斯聚点原理. 一对数 (x,y) 可以用 (x,y) 平面上以 x 和 y 为直角坐标的一个点来表示. 现在我们考虑这种点 $P(x,y)$ 的一个有界无穷集, 就是这样一个集合: 它包含无穷多个不同的点, 所有这些点都在平面上的一个有界的部分内, 所以 $|x|<C$ 并且 $|y|<C$, 其中 C 是一个常数. 聚点原理说的是: **每一个有界无穷点集** S **至少有一个聚点**. 这就是说, 存在一个点 $Q(\xi,\eta)$ 使得 Q 的每一个邻域内, 即每一个区域

$$(x-\xi)^2+(y-\eta)^2<\delta^2$$

(其中 δ 是任一个正数) 内, 都有 S 的无穷多个点. 由这聚点原理可以推知: 从无穷有界集中能够选出一序列不同的点 $P_1,P_2,P_3,\cdots$, 使它们收敛到一个极限 Q. 这些 P_i 可以用归纳法来构造, 这就是相继令 δ 为 $1,\frac{1}{2},\frac{1}{3},\cdots$; 在 S 内任取 P_1; 而如果 $P_1,\cdots,P_n$ 已经取定, 就在 S 内的无穷多个与 Q 的距离 $<\frac{1}{n+1}$ 的点中选一个不同于 Q 也不同于 $P_1,\cdots,P_n$ 的点作为 P_{n+1}.

多维聚点原理可以用第一卷 (第 105 页) 里相应的证明中用过的方法来类似地证明, 只要用矩形区域代替那里用的区间. 如果我们利用一维的这个原理可以得到一个更容易的证明. 首先注意到, 根据假设, 集合 S 的每一个点 $P(x,y)$ 的横坐标 x 满足不等式 $|x|<C$, 其次分两种情况来考虑: 或者有一个 $x=x_0$, 它是无穷多个点 P 的横坐标 (这些点 P 一个在另一个的正上方), 或者每一个 x 只属于有限个点 P. 对于第一种情况, 我们固定 x_0 而考虑 y 的无穷多个值使 (x_0,y) 属于我们的集合 S. 根据一维的聚点原理, y 的这些值有一个聚点, 因此我们可以找到 y 的一序列值 $y_1,y_2,\cdots$ 使 $y_n\to\eta_0$, 由此得到集合 S 的一序列点 (x_0,y_n), 收敛到极限点 (x_0,η_0), 它也

就是集合 S 的一个聚点. 对于第二种情况, 集合 S 的点的横坐标 x 一定有无穷多个不同的值, 我们可以取一序列横坐标 $x_1, x_2, \cdots$ 趋于一个极限 ξ. 对每一个 x_n, 取集合 S 的一个点 $P_n = (x_n, y_n)$. 这样 y_n 就构成一个无穷有界集, 因而可以取一个子序列 $y_{n_1}, y_{n_2}, \cdots$ 趋于一个极限 η. 相应的横坐标子序列 $x_{n_1}, x_{n_2}, \cdots$ 仍趋于极限 ξ; 于是点列 $P_{n_1}, P_{n_2}, \cdots$ 趋于极限点 (ξ, η). 这样, 就对于任何一种情况, 我们都可以找到集合 S 的一序列点趋于一个极限点, 从而这个定理就证明了.

b. 柯西收敛准则. 紧性

波尔札诺 - 魏尔斯特拉斯定理的一个推论是: **每一个有界的无穷点序列** $P_1, P_2, \cdots$ **必有一个收敛的子序列**. 如果这个序列包含无穷多个**不同**的元素, 那么它们构成一个无穷多个不同点的集合, 按魏尔斯特拉斯原理, 我们可以选出一个序列收敛到一个点 Q. 如果这个序列不包含无穷多个不同的元素, 那么它至少有一个元素重复了无穷多次; 于是存在一个点 Q 它在序列中出现无穷多次, 这些等于 Q 的元素构成的子序列就收敛到 Q.

一个重要的推论是**柯西收敛准则**:

平面上的一个点序列 $P_1, P_2, \cdots$ (**同样地**, n **维欧氏空间的一个点序列**)**收敛到一个极限当且仅当对每一个** $\epsilon > 0$ **存在一个数** $N = N(\epsilon)$ **使得当** n **和** m **大于** N **时** P_n **和** P_m **之间的距离小于** ε.

证明的步骤与第一卷 (106 页) 中对实数序列给出的相应的证明完全一样. 我们立刻看出, 一个序列要是满足柯西条件它就是有界的; 所以根据前面的定理, 它就包含一个收敛子序列, 以一点 Q 为极限, 于是立刻可以推导出这整个序列收敛到 Q.

我们说过, 平面上的一个点集 S 叫做**闭的**, 如果 S 的一切边界点属于 S. 闭集 S 中的每一个收敛的点列的极限 Q 也是 S 的点

(见第 9 页). 因为每个有界无穷序列包含一个收敛子序列, 所以我们得知: 平面上的一个有界闭集 S 中的每一个无穷序列都包含一个子序列收敛到 S 的一个点. 一般地, 我们称一个集合 S 为紧的 [1], 如果 S 的元素构成的每一个序列都包含一个收敛的子序列, 其极限在 S 内. 所以, 平面(或 n 维欧氏空间)上的闭的有界点集都是紧的. 读者可以容易地验证逆命题: 平面上每一个紧的点集都是闭的并且是有界的. 今后我们将经常在提到闭的有界集时简单地称之为紧集(compact set).

c. 海涅 - 波瑞耳覆盖定理

波尔札诺 - 魏尔斯特拉斯原理的一个突出的推论是海涅 - 波瑞耳定理 (Heine-Borel theorem):

设 S 是一个紧集(即闭的并且有界), Σ 是一类(无穷多个)开集, 它覆盖了 S, 就是说 S 的每一个点至少属于 Σ 中的一个开集那么我们可以在 Σ 中找到有限个开集, 它们已经覆盖了 S.

作为一个例子, 我们考虑无穷点集 S, 它是由 x 轴上的点 $P_n = \left(\frac{1}{n}, 0\right)$, $n = 1, 2, \cdots$, 和原点 $P_0 = (0, 0)$ 组成的. 这是一个闭集. 对 $n = 1, 2, \cdots$, 令 S_n 为以 P_n 为中心, $1/3n^2$ 为半径的开圆:

$$\sqrt{\left(x - \frac{1}{n}\right)^2 + y^2} < \frac{1}{3n^2},$$

再令 S_0 表示开圆

$$\sqrt{x^2 + y^2} < \frac{1}{100}.$$

显然这无穷多个集合 $S_0, S_1, S_2, \cdots$ 覆盖了 S. 根据海涅 - 波瑞耳定理, 我们可以在其中选出有限个就足以覆盖 S, 例如选取 $S_0, S_1, \cdots, S_{100}$. 在这里我们立刻看到 S 是闭的假设的重要性. 设

1) 有时更精确地说 "序列列紧".

集合 T 是由点 $P_1, P_2, \cdots$ 组成，没有 P_0，那么它被 $S_1, S_2, \cdots$ 所覆盖，但这时因为每一个 S_i 只包含 T 的一个点，所以不可能选出有限个 S_i 来覆盖 T.

我们用反证法来证明海涅 - 波瑞耳定理. 设这定理不成立. 集合 S 既然有界，它就位于一个方块 Q 内. 我们把 Q 分成四个相等的方块. 这四个方块中至少有一个方块，加上它的边界后，所包含的那部分 S 不能被 Σ 中的有限个开集合所覆盖. 因为如果 S 的这四部分都能被有限个开集合覆盖，那么 S 本身也就可以被有限个开集合覆盖了. 我们把 Q 的这一部分叫做 Q_1. 我们现在把 Q_1 也分成四个相等的部分. 经过同样的推理可取出 Q_1 的四个部分之一作为 Q_2，这 Q_2 和它边界上的 S 的点不能被 Σ 中有限个开集所覆盖. 继续这样进行下去，我们就得到一无穷序列方块 $Q_1, Q_2, Q_3, \cdots$，每一个包含在它的前一个中，它们的面积收缩为 0，并且每一个 Q_n 的闭包中属于 S 的点不可能被 Σ 中有限个开集所覆盖. 显然，对每一个 n，我们可以在 Q_n 内或它的边界上找到一个属于 S 的点 P_n. 于是 $P_1, P_2, \cdots$ 是 S 中的一序列点. 因为 S 是有界的，这个点列也是有界的，因而必定有一个收敛的子序列收敛到某点 A. 又因为 S 是闭的，这 A 也是 S 的一个点，因而被包含在 Σ 中的某一个开集 Ω 之内. 从而有 A 的一整个邻域在该开集 Ω 内，譬如说是由与 A 的距离小于 ε 的点所组成的邻域. 我们可以取 n 充分大，使得 P_n 到 A 的距离小于 $\frac{\varepsilon}{2}$，并且 Q_n 的直径也小于 $\frac{\varepsilon}{2}$. 这样，这整个方块 Q_n 就包含在 A 的 ε 邻域内，从而包含在 Ω 内了. 我们看到 Σ 中的这一个集合 Ω 就包含了整个方块 Q_n 和它的边界，而这却与关于方块 Q_n 的假设矛盾. 证完.

d. 海涅 - 波瑞耳定理在开集所包含的闭集上的应用

令 R 是平面上的一个开集[1)]，根据定义， R 的每一个点 P 有一个邻域完全在 R 内. 对于靠近 R 的边界的点 P，邻域非常小. 值得注意的是当 P 属于 R 的一个闭子集 S 时，我们可以找到一个

1) 这一段里说到的每一件事都同样地适用于高维，只要我们把"圆"改成"球".

一致的大小 ε, 使 P 的 ε 邻域全包含在 R 内: 如果一个闭的有界集 S 包含在开集 R 内, 则存在一个正数 ε 使得 S 的每一个点 P 的 ε 邻域都包含在 R 内. 换句话说, 不在 R 内的点到 S 的点的距离至少是 ε [1].

为了证明, 我们利用 R 是开的这一假设. 对于 R 的每一个点 P 存在一个以 P 为心的圆, 包含在 R 内. 这个圆的半径叫 r, 它依赖于 P, 即 $r = r(P)$. 现在我们对 S 内的任一点 P 取以 P 为心半径是 $\frac{1}{2}r(P)$ 的圆. 根据海涅 - 波瑞耳定理可以找到有限个这种圆覆盖住这个紧集 S. 因此我们找到 S 内的有限个点 $P_1, P_2, \cdots, P_n$, 使 S 的每一个点 P 包含在某一个以 P_k 为中心、以 $\frac{1}{2}r(P_k)$ 为半径的圆内 $(k = 1, \cdots, n)$. 令 ε 是正数 $\frac{1}{2}r(P_1), \cdots, \frac{1}{2}r(P_n)$ 中的最小者. 那么, 对 S 的每一个点 P, 由于 P 在某一个以 P_k 为中心、以 $\frac{1}{2}r(P_k)$ 为半径的圆内, 所以 P 的 ε 邻域在 R 内. 这是因为以 P_k 为中心、以 $r(P_k)$ 为半径的同心圆 D 是在 R 内, 而由于

$$\overline{PP_k} < \frac{1}{2}r(P_k) \text{ 和 } \varepsilon \leq \frac{1}{2}r(P_k),$$

圆 D 就包含了以 P 为中心、以 ε 为半径的圆. 这就证明了以 P 为中心、以 ε 为半径的圆在 R 内.

作为例子, 考虑开集 R 内的曲线 S. 这样一条曲线是由这样的点 $P = (x, y)$ 所成的集合, 这些点可以表成这样的形式

$$x = \varphi(t), \quad y = \psi(t),$$

其中 φ, ψ 是两个连续函数, 参量 t 在闭区间 $0 \leq t \leq 1$ 上 [2]. 这样一条曲线 S 是一个闭的点集. 事实上, 设 $P_1, P_2, \cdots$ 是 S 的一

---

1) S 有界是必要的. 例如, R 是开的半平面 $y > 0$. S 是闭集包含 (x, y) 平面上的这种点: $y \geq \frac{1}{x}$, $x > 0$; 那么 R 的边界就可以任意接近 S 的点.

2) 这曲线不要求是简单的; 也就是, 不同的 t 可以对应相同的点 P. 这一对函数定义了一条 "路径", S 是路径的支集.

个点序列，收敛到一点 P. 我们考虑相应的参量值 $t_1, t_2, \cdots$, 它们都在闭区间 $a \le t \le b$ 上. 因为闭的有界区间是紧的，所以有 t_n 的一个子序列收敛到区间的一个值 t. 但因为 φ 和 ψ 是连续的，所以相应的 P_n 收敛到 S 上的点 $Q = (x(t), y(t))$. 这样就有了序列 $P_1, P_2, \cdots$ 的一个子序列收敛到 S 的一点 Q. 然而整个序列是收敛到点 P 的，所以 $P = Q$, 因而 P 在 S 上. 这说明 S 包含了所有 S 的点序列的极限，因而 S 是闭的.

如果曲线在开集 R 内，我们就可以找到一个正数 ϵ 使得所有以 S 的点为中心、以 ϵ 为半径的圆都在 R 内. 因为 φ 和 ψ 是连续的，因而是一致连续的，我们可以找到一个正数 δ, 使得当两个参量之差小于 δ 时，它们对应的 S 上的点的距离小于 ϵ. 我们用点 $t_1, \cdots, t_{n-1}$ 分割参变量区间

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b,$$

其中每个子区间的长小于 δ. 令 $P_0, P_1, \cdots, P_n$ 是 S 上的相应的点. 那么 P_{i+1} 总是在 P_i 的 ϵ 圆域内. 从而联接 P_i 和 P_{i+1} 的直线段也完全在以 P_i 为中心、以 ϵ 为半径的圆内，也就包含在 R 内. 如果我们用直线段依次联接 P_i, 就得到一条**折线**完全在 R 内，并与连续曲线 S 有相同的端点 P_0 和 P_n. 我们可以把这结果叙述如下：

如果开集 R 的两个点可以用 R 内的一条曲线联接起来，则它们也可以用 R 内的一条折线联接起来.

A.2 连续函数的基本性质

对于在一个闭的有界集 S 上有定义并且连续的函数 f, 我们可以叙述下面两个基本定理：

函数 f 在 S 上取到最大值和最小值.

函数 f 在 S 上一致连续.

这两个定理的证明和一元函数中相应的证明一样 (见第一卷第 110—112 页), 不需要重复了.

第二个定理也可以作为海涅 - 波瑞耳定理的一个直接推论而得到. 指定一个 $\varepsilon > 0$. 若 f 在 S 的每一点连续, 则对 S 的每一个点 P 存在一个以 P 为圆心、以某个 $\delta = \delta(P)$ 为半径的 δ 邻域, 使得 S 内的任意点 Q 只要它在这个邻域内就有 $|f(Q) - f(P)| < \varepsilon/2$. 现在对 S 内的每一点 P 取一个半径为 $\frac{1}{2}\delta(P)$ 的邻域 Ω_p. 显然这种 Ω_p 覆盖了 S. 我们可以从中选取有限个 (中心为 $P_1, \cdots, P_n$) 也覆盖了 S. 令 Δ 是数 $\frac{1}{2}\delta(P_1), \cdots, \frac{1}{2}\delta(P_n)$ 中的最小者. 如果 P 和 Q 是 S 的任意两个点, 其间距离小于 Δ, 则因点 P 到某个点 P_k 的距离小于 $1/2\delta(P_k)$ $(k = 1, \cdots, n)$, 而 $\Delta \leq \frac{1}{2}\delta(P_k)$, 我们就看到, P 和 Q 都在点 P_k 的 $\delta(P_k)$ 邻域内. 所以

$$|f(P) - f(P_k)| < \frac{1}{2}\varepsilon, \qquad |f(Q) - f(P_k)| < \frac{1}{2}\varepsilon,$$

从而

$$|f(P) - f(Q)| < \varepsilon.$$

因为 Δ 不依赖于 P 和 Q 的特定的位置, 这就证明了 f 的一致连续性.

A.3 点集论的基本概念

a. 集合与子集合

在关于点的集合的更为复杂的讨论中 (特别在积分理论中), 使用一些标准的集合运算记号是方便的. 我们有兴趣的集合经常是数的、点的、函数的集合, 或者是这些类型的集合的集合. 例如平面上的一个 "圆" 是这样的点 (x, y) 所成的集合, 他们对固定的 x_0, y_0, r 满足

$$(x - x_0)^2 + (y - y_0)^2 < r^2.$$

一个集合的集合 (集合族) 的例子是：所有包含原点的圆，也就是上面所说的那样的适合条件 $x_0^2+y_0^2<r^2$ 的一切圆.

我们不再把 集合 的基本概念化为更加基本的概念，也不去分析涉及这个概念的逻辑困难. 对我们来说，一个集合 S 被认为是确定的，只要对于每一个对象 α 都确有下面两个情形之一成立：(1) α 属于 S; (2) α 不属于 S. 对情形 (1), 我们也说 α 是 S 的 一个元素, 或说 α 包含在 S 内; 用符号表示成 $\alpha\in S$; 1)

对情形 (2), 记作 $\alpha\notin S$.

例如, 如果 S 是由不等式 $x^2+y^2<r^2$ 给定的圆, 那么 $\alpha\in S$ 的意思是：α 是平面上的一个点, 它的坐标 x,y 具有性质 $x^2+y^2<r^2$. 一般地说，一个集合 S 的元素可以被某些共同的 性质 所刻画 (例如，用属于 S 这个性质). 我们把具有性质 $A,B,\cdots$ 的元素 α 所组成的集合 S 记作

$$S=\{\alpha:\alpha \text{ 有性质 } A,B\cdots\}.$$

例如，以 (x_0,y_0) 为中心、以 r 为半径的圆 S 可以记作

$$S=\left\{(x,y):x,y=\text{实数};(x-x_0)^2+(y-y_0)^2<r^2\right\}.$$

集合

$$S=\{n:n=\text{整数};2<n<5\}$$

是由 $n=3$, 和 $n=4$ 两个元素组成的.

为了许多目的，引进 “空”(或 “零”) 集及其特殊记号 $\varnothing$是方便的. 空集合没有元素：对一切 $\alpha,\alpha\notin\varnothing$. 例如，一个中心在原点半径是 0 的开圆就是 $\varnothing$:

$$\{(x,y):x,y=\text{实数};x^2+y^2<0\}=\varnothing.$$

1) 符号 $\in$ 切勿与希腊字母 ε 混淆.

两个集合 S 和 T 是相等的，只要它们有共同的元素，而不管用在它们定义里的描写或性质有所不同．也就是： $S=T$ 的意思是 $x\in S$ 当且仅当 $x\in T$.

我们说集合 S 是集合 T 的一个**子集**（"S 包含在 T 内"），如果 T 包含一切包含在 S 里的元素．这就是说， $\alpha\in S$ 蕴涵 $\alpha\in T$. 我们用符号记作 $S\subset T$, 或较少地记作 $T\supset S$. 例如，如果 S 是以原点为中心、以 1 为半径的圆， T 是以 (1,1) 为心、以 4 为半径的圆，就有 $S\subset T$. 类似地，对一切集合 S 都有 $\emptyset\subset S$ 和 $S\subset S$.

当然，符号 $\subset$ 和 $\supset$ 可以选用，它们就像算术里的符号 $<$ 和 $>$ (或者更精确地说是像 $\le$ 和 $\ge$)．它们在下面的基本性质上和后面这两个符号一样

$$S\subset T\text{和}T\subset S\text{蕴涵}S=T,$$

$$S\subset T\text{和}T\subset R\text{蕴涵}S\subset R\ ^{1)}.$$

在集合的"包含于"符号和数的顺序符号之间基本的不同是：对于实数我们总有 $x\le y$ 或 $y\le x$, 而对于集合则可能有这样的情况：命题 $S\subset T$ 或 $T\subset S$ 都不成立．符号 $\subset$ 只是在集合之间定义了"偏"序；可能两个集合中的任一个都不包含另一个．

b. 集合的并与交

近几十年来大量的逻辑符号在数学里被广泛地采用，所以现在习惯于把很多数学定理完全用符号来表示，而不用平常的文字或语句 [2)]．从最古的时候起，使用适当的符号性记号就是很重要的；事

1) 这是逻辑里的普通三段论法：如果一切有性质 A 的对象都有性质 B, 又一切有性质 B 的对象都有性质 C, 那么一切有性质 A 的对象都有性质 C.

2) 经常用的符号的例子如下：

$\{x_1,x_2,\cdots,x_n\}$: 这集合的元素明确地是 $x_1,\cdots,x_n$.

$S\times T$: 有序的数对 (a,b) 的集合，其中 $a\in S, b\in T$ (集合 S,T 的"笛卡儿乘积")

$\rightarrow$:"蕴涵".

$\exists x$:"存在一个 x".

$\forall x$:"对一切 x".

实上，有这种例子，正是由于缺乏适当的符号使得某些领域的进展放慢达几个世纪，可以说代数在古代就是这样．但另一方面，过分集中使用符号的结果可能给一个试图把这种“脱水”形式的知识与他的日常经验联系起来的读者带来很大的困难．不是逻辑或数学基础专著的一些书籍的作者们，在使用逻辑缩写上，是按照他们自己的爱好和所考虑的特定问题的需要而采取折衷的方案的．

还有两个集合论的符号，就是集合运算“并”和“交”的符号，我们将在本书的后面看到它们几乎是必不可缺少的．给定两个集合 S 和 T，我们用 $S\cup T$ 表示这两个集合的“并”，就是“或”在 S 内、“或”在 T 内的元素组成的集合：

$$S\cup T=\{a:a\in S\ \text{或}\ a\in T\}^{1)}.$$

类似地，S 和 T 的“交”$S\cap T$ 定义为属于 S 和 T 两者的元素的集合：

$$S\cap T=\{a:a\in S\ \text{且}\ a\in T\}.$$

例如，S 和 T 是实数轴上的区间

$$S=\{x:3<x<5\},\qquad T=\{x:4\le x<6\},$$

则

$$S\cup T=\{x:3<x<6\},$$
$$S\cap T=\{x:4\le x<5\}.$$

运算 $\cup$ 和 $\cap$ 适用于任意两个集合 S 和 T，只要我们引用空集的符号，当 S 和 T **不相交**时可以记作

$$S\cap T=\varnothing,$$

也就是没有公共元素．注意，对任意 S 都有

$$S\cup\varnothing=S,\quad S\cap\varnothing=\varnothing.$$

---

1) 这里“或”字像拉丁文 vel 一样，是不相排斥的．$S\cup T$ 的元素是至少属于集合 S,T 两者之一的，也可以属于两者．

运算 $\cup$ 有很多性质和加法一样. 特别地, 当 S 和 T 是不相交的集合 —— 就是这两个集合没有公共元素 —— 并且都只有有限个元素时, $S\cup T$ 内的元素的个数恰好是 S 和 T 的元素的个数之和. 可是, 一般说来, 并的逆运算不唯一. 只有当 S 和 T 不相交并且 $S\subset R$ 时, 方程

$$S\cup T=R$$

才有唯一的解 T. 对于不相交的集合 S 和 T, 它们的并常记作 $S+T$, 当 $S\subset R$, 方程 $S+T=R$ 的解是 $R-S$ ("S 关于 R 的余"). 我们将要更为普遍地对任意集合 R,S 使用符号 $R-S$, 它表示 R 内的不属于 S 的元素所组成的集合. 于是 $S+(R-S)=R\cup S$.

n 个集合 $S_1,\cdots,S_n$ 的并定义为至少属于集合 $S_1,\cdots,S_n$ 之一的那些元素所组成的集合, 有几种记法

$$\begin{aligned}&\{a: a\in S_1 \text{ 或 } a\in S_2 \text{ 或 } \cdots \text{ 或 } a\in S_n\}\\ =&S_1\cup S_2\cup\cdots\cup S_n\\ =&\bigcup_{k=1}^{n} S_k,\end{aligned}$$

类似于和号. 类似地, 集合 $S_1,\cdots,S_n$ 的交定义为所有这些集合的公共元素所组成的集合, 记作

$$\begin{aligned}&\{a: a\in S_1 \text{ 且 } a\in S_2 \text{ 且 } \cdots \text{ 且 } a\in S_n\}\\ =&S_1\cap S_2\cap\cdots\cap S_n\\ =&\bigcap_{k=1}^{n} S_k.\end{aligned}$$

我们可以同样地作无穷多个集合 $S_1,S_2,\cdots,S_n,\cdots$ 的并和交, 分别记作

$$\bigcup_{k=1}^{\infty} S_k=\{a: a\in S_n \text{ 对某个 } n\},$$

$$\bigcap_{k=1}^{\infty} S_k=\{a: a\in S_n \text{ 对一切 } n\}.$$

举例，如果 S_n 是实数 $x<n$ 的集合

$$S_n=\{x : x \text{ 实数}, x<n\},$$

我们有

$$\bigcup_{k=1}^{\infty} S_k=\{x : x \text{ 实数}\},$$
$$\bigcap_{k=1}^{\infty} S_k=\{x : x \text{ 实数}, x<1\}.$$

事实上，并和交可以对集合 S 的任意广大的族 F 来做，甚至 F 中不同的集合 S 不是，或不能，用下标 $n, n=1,2,3,\cdots$ 来区别，也一样进行．我们记

$$\bigcup_{S\in F} S=\{a : a\in S \text{ 对某 } S,\ S\in F\},$$
$$\bigcap_{S\in F} S=\{a : a\in S \text{ 对一切 } S,\ S\in F\}.$$

例如 (x,y) 平面上包含点 $(1,0)$ 但不包含点 $(-1,0)$ 的一切圆的并集是适合条件 $y\neq 0$ 或 $y=0$ 且 $x>-1$ 的全部点 (x,y) 所成的集合．而这一族圆的交则只包含 $(1,0)$ 这一个点．

c. 应用于平面上的点集

前面的某些结果和定义（见第 6—8 页）可以用上面所引进的表示方法重新改写得更为简洁．给定平面上的一个点集 S，我们得到全平面 π 的一个分解，分为三个不相交的集合：S 的内点集合 S^0，S 的边界点集合 ∂S，S 的外点集合 S_e，于是

$$\pi=S^0\cup\partial S\cup S_e,$$

或者更准确地写成

$$\pi=S^0+\partial S+S_e.$$

因为这些集合是不相交的：

$$S^0\cap\partial S=\partial S\cap S_e=S_e\cap S^0=\emptyset,$$

这里

$$S^0 \subset S \subset S^0 + \partial S.$$

如下定义的集合 $\overline{S}$ 是 S 的**闭包**:

$$\overline{S} = S^0 + \partial S = S \cup \partial S. \tag{1}$$

对于开集 S 我们有 $S^0 = S$; 对于闭集有 $\overline{S} = S$.

读者可以作为练习验证下面的性质:

$\overline{\partial S} = \partial S$ ("一个集合的边界总是闭的"),

$\overline{\overline{S}} = \overline{S}$ ("一个集合的闭包总是闭的"),

$(S^0)^0 = S^0$, $(S_e)^0 = S_e$ ("集合 S^0 和 S_e 是开的").

$$S^0 \cup T^0 \subset (S \cup T)^0, \quad \overline{S \cup T} \subset \overline{S} \cup \overline{T}. \tag{2a}$$

$$\partial(S \cup T) \subset \partial S \cup \partial T. \tag{2b}$$

开集的并是开的.

有限个闭集的并是闭的.

有限个开集的交是开的.

闭集的交是闭的.

这最后几个命题指出了在 "开" 和 "闭","并" 和 "交" 的概念之间的一种对称性 ("对偶性"). 如果我们引进集合 S 的**余集** $C(S)$, 也就是平面 π 上不属于 S 的点所形成的集合 [1]

$$C(S) = \{P : P \in \pi, P \notin S\} = \pi - S,$$

那么上面的对称性就更加明显. 我们有

$$C(S^0) = \overline{S}_e, \quad \partial C(S) = \partial S, \quad C(S_e) = \overline{S}_0.$$

---

1) 对三维空间 Σ 的点集 S, S 的余集定义为 $\Sigma - S$, 它是 Σ 中不属于 S 的点所组成的集合.

如果 S 是开的，则 $C(S)$ 是闭的，反过来也对. 几个集合的交的余集是它们的余集的并.

使用这种表示法，海涅 - 波瑞耳定理得到特别简单的形式. "集合族 F 覆盖一个集合 S" 的意思是，S 包含在 F 的集合的并之中. 于是这个定理可简单地叙述成：

如果 F 是平面上的开集族，又如果 S 是一个有界闭集，而有

$$S \subset \bigcup_{T \in F} T,$$

那么可以找到有限个集合 $T_1, T_2, \cdots, T_n \in F$ 使得

$$S \subset \bigcup_{k=1}^{n} T_k.$$

A.4 齐次函数

出现在分析及其应用里的最简单的齐次函数是型或多变量的齐次多项式 (见第 14 页). 我们称形如 $ax + by$ 的函数为 x 和 y 的一阶齐次函数，称形如 $ax^2 + bxy + cy^2$ 的函数为二阶齐次函数，而一般地称一个 x 和 y 的多项式(或者再多几个变量)为 h 阶的齐次函数，如果在每一项里自变量的指数的和都等于 h. 就是说，如果这些项 (去掉常系数) 具有这样的形式 $x^h, x^{h-1}y, x^{h-2}y^2, \cdots, y^h$. 齐次多项式具有以下性质：方程

$$f(tx, ty) = t^h f(x, y)$$

对每一个 t 的值都成立. 更为一般地，我们说一个函数 $f(x, y, \cdots)$ 是 h 阶齐次的，如果它满足方程

$$f(tx, ty, \cdots) = t^h f(x, y, \cdots).$$

不是多项式的齐次函数的例子有

$$\tan\left(\frac{y}{x}\right) \quad (h = 0),$$

$$x^2 \sin\frac{x}{y} + y\sqrt{x^2+y^2}\log\frac{x+y}{y} \quad (h=2).$$

另外的例子是分别以 x, y, z 和 u, v, w 为分量的两个向量之间的夹角的余弦

$$\frac{xu+yv+zw}{\sqrt{x^2+y^2+z^2}\cdot\sqrt{u^2+v^2+w^2}} \quad (h=0).$$

以 x, y, z 为分量的向量长度

$$\sqrt{x^2+y^2+z^2}$$

是一个一阶 **正齐次** 函数的例子；所谓正齐次函数是指定义齐次函数的那个方程式对它来说除非 t 是正的或 0, 否则方程不成立.

齐次函数和它的微商满足欧拉偏微分方程:

$$xf_x + yf_y + zf_z + \cdots = hf(x, y, z, \cdots).$$

为了证明这个结论，我们在方程 $f(tx, ty, \cdots) = t^h f(x, y, \cdots)$ 的两端对 t 求微商，这是允许的，因为方程对 t 是恒等的. 对左端的函数应用微商的锁链法则得到

$$xf_x(tx, ty, \cdots) + yf_y(tx, ty, \cdots) + \cdots = ht^{h-1}f(x, y, \cdots).$$

令 $t = 1$ 就得到结论.

反过来，不难证明函数 $f(x, y, \cdots)$ 的齐次性是欧拉关系的推论，所以 **欧拉关系是函数有齐次性的必要充分条件**. 一个函数是 h 阶齐次函数这个事实，也可以表示为用 x^h 除函数值后只依赖于 $\frac{y}{x}, \frac{z}{x}, \cdots$. 所以只要证明由欧拉关系可以推出，在引进新变量

$$\xi = x, \quad \eta = \frac{y}{x}, \quad \zeta = \frac{z}{x}, \cdots$$

之后函数

$$\frac{1}{x^h}f(x, y, z, \cdots) = \frac{1}{\xi^h}f(\xi, \eta\xi, \zeta\xi, \cdots) = g(\xi, \eta, \zeta, \cdots)$$

不再依赖于变量 ξ (即方程 $g_\xi = 0$ 是恒等式). 为了证明这一点, 们用锁链法则写出

$$g_\xi = (f_x + \eta f_y + \cdots)\frac{1}{\xi^h} - \frac{h}{\xi^{h+1}}f$$
$$= (xf_x + yf_y + \cdots)\frac{1}{x^{h+1}} - \frac{h}{x^{h+1}}f.$$

再根据欧拉关系, 右边为零, 我们的结论就证明了.

最后这个结论也可以用更为巧妙、但不直接的方法来证明. 我们来证: 由欧拉关系可以推导出函数

$$g(t) = t^h f(x, y, \cdots) - f(tx, ty, \cdots)$$

对一切 t 取零值. 显然有 $g(1) = 0$. 又

$$g'(t) = ht^{h-1}f(x, y, \cdots) - xf_x(tx, ty, \cdots) - yf_y(tx, ty, \cdots) \cdots.$$

对变量 $tx, ty, \cdots$ 应用欧拉关系, 我们有

$$xf_x(tx, ty, \cdots) + yf_y(tx, ty, \cdots) + \cdots = \frac{h}{t}f(tx, ty, \cdots),$$

所以 $g(t)$ 满足微分方程

$$g'(t) = g(t) \cdot \frac{h}{t}.$$

如果我们令 $g(t) = r(t)t^h$, 就得到

$$g'(t) = \frac{h}{t}g(t) + t^h r'(t),$$

所以 $r(t)$ 满足微分方程

$$t^h r'(t) = 0.$$

它有唯一解 r = 常数 = c. 因为 $t = 1$ 有 $r(t) = 0$, 所以常数 c 是 0 也就是, 对于一切 t 的值确有 $g(t) = 0$. 证完.

第二章 向量、矩阵与线性变换

在第一卷第四章中，我们讨论了二维向量. 高维的几何概念使向量的应用甚至更为重要. 向量适合于简明地表示众多复杂的方程，因而在一定程度上清楚地显示出那些不依赖于坐标系的特殊选择的特征.

2.1 向量的运算

a. 向量的定义

在 n 维空间中，我们引进向量，作为可以相加并且与数量相乘的实体. 明确地说，一个向量 $\mathbf{A}$ 是一组有确定顺序的 n 个实数 [1)] $a_1, a_2, \cdots, a_n$:

$$\mathbf{A} = (a_1, a_2, \cdots, a_n).$$

(我们总是用黑体字表示向量.) 数量 $a_1, a_2, \cdots, a_n$ 称为向量 $\mathbf{A}$ 的**分量**. 两个向量 $\mathbf{A} = (a_1, a_2, \cdots, a_n)$ 与 $\mathbf{B} = (b_1, b_2, \cdots, b_n)$ 是相等的，当且仅当它们有相同的分量.

任意两个向量 $\mathbf{A} = (a_1, a_2, \cdots, a_n)$ 与 $\mathbf{B} = (b_1, b_2, \cdots, b_n)$ 的和定义为

$$\mathbf{A} + \mathbf{B} = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n); \tag{1a}$$

---

1) 对我们的目的来说，只需考虑实数分量就够了，虽然在别的教科书中也使用其他数域上的向量.

我们定义向量 $\mathbf{A}$ 与数量 (即实数) λ 的乘积为

$$\lambda\mathbf{A} = (\lambda a_1, \lambda a_2, \cdots, \lambda a_n)\ ^{1)}. \tag{1b}$$

更一般地，我们能够从任意有限个向量 $\mathbf{A} = (a_1, a_2, \cdots, a_n)$, $\mathbf{B} = (b_1, b_2, \cdots, b_n), \cdots, \mathbf{D} = (d_1, d_2, \cdots, d_n)$ 与相等个数的数量 $\lambda, \mu, \cdots, \nu$ 作成**线性组合**

$$\begin{aligned}\lambda\mathbf{A} + \mu\mathbf{B} + \cdots + \nu\mathbf{D} = (&\lambda a_1 + \mu b_1 + \cdots + \nu d_1,\\ &\lambda a_2 + \mu b_2 + \cdots + \nu d_2, \cdots,\\ &\lambda a_n + \mu b_n + \cdots + \nu d_n).\end{aligned}$$

特别地，任何一个向量 $\mathbf{A} = (a_1, a_2, \cdots, a_n)$ 都能够表示成 n 个 "坐标向量"

$$\begin{gathered}\mathbf{E}_1 = (1, 0, \cdots, 0), \quad \mathbf{E}_2 = (0, 1, 0, \cdots, 0), \cdots,\\ \mathbf{E}_n = (0, \cdots, 0, 1)\end{gathered} \tag{2a}$$

的一个线性组合. 显然

$$\mathbf{A} = a_1\mathbf{E}_1 + a_2\mathbf{E}_2 + \cdots + a_n\mathbf{E}_n. \tag{2b}$$

我们用符号 $\mathbf{O}$ 表示 "零向量", 它的所有分量都是零: $\mathbf{O} = (0, 0, \cdots, 0)$. 我们把向量 $(-1)\mathbf{A} = (-a_1, -a_2, \cdots, -a_n)$ 记作 $-\mathbf{A}$.

从这些定义容易推出, 向量的和以及同数量的乘积满足全部通常的代数定律, 只要它们有意义即可 [2]. 由有限个适当选定的函数

1) 向量不同于其他能用 n 个有序实数描写的对象 (例如 n 维欧几里得空间中的点或 $n+1$ 维球面上的点), 就在于这个事实, 即向量允许 "线性运算" $\mathbf{A}+\mathbf{B}$ 和 $\lambda\mathbf{A}$. 用点的坐标类似地定义点的加法将没有几何意义, 至少, 离开所用的特定的坐标系, 它将是无意义的. 后面, 向量将表示成点对 (见第 135 页).

2) 这些定律如下:

(1) $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$, $\mathbf{A}+(\mathbf{B}+\mathbf{C}) = (\mathbf{A}+\mathbf{B})+\mathbf{C}$

(2) $\lambda(\mathbf{A}+\mathbf{B}) = \lambda\mathbf{A}+\lambda\mathbf{B}$, $(\lambda+\mu)\mathbf{A} = \lambda\mathbf{A}+\mu\mathbf{A}$, $(\lambda\mu)\mathbf{A} = \lambda(\mu\mathbf{A})$

(3) 存在唯一的一个元素 $\mathbf{O}$, 使得对于每一个向量 $\mathbf{A}$ 都有 $\mathbf{A}+\mathbf{O} = \mathbf{A}$

(4) 对于给定的 $\mathbf{A}$, 存在唯一的一个元素 $-\mathbf{A}$, 使得 $\mathbf{A}+(-\mathbf{A}) = \mathbf{O}$

(5) 对于所有的 $\mathbf{A}$ 都有: $0\mathbf{A} = \mathbf{O}$, $1\mathbf{A} = \mathbf{A}$.

一般来说, 一个集合, 对其元素定义了加法以及数乘法, 并且遵守这些规律, 就叫做向量空间.

的线性组合而成的函数，提供了便于用向量表示的对象的例子. 例如，变量 x 的次数 $\leq n$ 的一般多项式

$$P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

就能够用 $n+1$ 维空间中的一个单独的向量 $\mathbf{A} = (a_0, a_1, \cdots, a_n)$ 来表示. 于是，向量的加法与数乘法就对应于对多项式所施行的相同的运算. 类似地，一般的 n 阶三角多项式

$$f(x) = \frac{1}{2}a_0 + \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx)$$

(见第一卷第 643 页) 能够用 $2n+1$ 维空间中的向量 $(a_0, a_1, \cdots, a_n, b_1, \cdots, b_n)$ 来表示. 一般的三个变量的线性齐次函数

$$u = a_1x_1 + a_2x_2 + a_3x_3$$

可用三维空间中的向量 (a_1, a_2, a_3) 表示; 一般的三个变量的二次型

$$u = a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + 2a_4x_2x_3 + 2a_5x_3x_1 + 2a_6x_1x_2$$

可用六维空间中的向量 $(a_1, a_2, \cdots, a_6)$ 表示.

b. 向量的几何表示

正如平面情形一样，能够把 n 维空间中的向量几何地看作这空间的某个映射，即平移或平行位移. 向量 $\mathbf{A} = (a_1, a_2, \cdots, a_n)$ 可以用来描写 n 维欧几里得空间 R^n 的平移，该平移将任意点 $P = (x_1, x_2, \cdots, x_n)$ 映射到 $P' = (x_1', x_2', \cdots, x_n')$，它们的坐标之间有如下关系 [1)]

$$x_1' = x_1 + a_1, \quad x_2' = x_2 + a_2, \cdots, \quad x_n' = x_n + a_n. \tag{3a}$$

1) 我们约定，两个点 P 和 P' 都同在空间 R^n 中，并且它们坐标是关于同一坐标系而取的.

假如对于某一个点 $P=(x_1,x_2,\cdots,x_n)$，我们给出了像点 $P'=(x'_1,x'_2,\cdots,x'_n)$，则这个平移或其对应的向量 $\mathbf{A}$ 就唯一地被确定了；由 (3a) 显然有

$$\mathbf{A}=(x'_1-x_1,x'_2-x_2,\cdots,x'_n-x_n). \tag{3b}$$

我们将用 $\mathbf{A}=\overrightarrow{PP'}$ 表示这个平移，并且说向量 $\mathbf{A}$ 是由这有序的点偶 P 和 P' 来表示. 在这种表示法中，我们称 P 为起点，P' 为终点. 在绘图中，常用从 P 到 P' 的一个箭头来表示向量 $\mathbf{A}=\overrightarrow{PP'}$. 同一个向量 $\mathbf{A}$，可以被很多对点偶 P 和 P' 构成的有向线段 $\overrightarrow{PP'}$ 来表示. 起点 P 完全是任意的，因为由 $\mathbf{A}$ 定义的变换可以作用在任何一个点上，而随之确定了一个像点 P' [1]. 零向量 $\mathbf{O}$ 对应着"恒等映射"，其中每一点都变到它自身：$\mathbf{O}=\overrightarrow{PP}$.

和平面情形 (第一卷第 431 页) 一样，两个向量 $\mathbf{A}=(a_1,a_2,\cdots,a_n)$，$\mathbf{B}=(b_1,b_2,\cdots,b_n)$ 的和产生它们对应映射的符号乘积. 如果 $\mathbf{A}$ 将点 $P=(x_1,x_2,\cdots,x_n)$ 变到点 $P'=(x'_1,x'_2,\cdots,x'_n)$，$\mathbf{B}$ 将点 P' 变到 $P''=(x''_1,x''_2,\cdots,x''_n)$，则 $\mathbf{C}=\mathbf{A}+\mathbf{B}$ 就对应着将 P 变到 P'' 的平移变换，因为

$$x''_i=x'_i+b_i=(x_i+a_i)+b_i=x_i+(a_i+b_i)$$

对于 $i=1,2,\cdots,n$ 成立. 用向量的记号就有

$$\mathbf{A}+\mathbf{B}=\overrightarrow{PP'}+\overrightarrow{PP'''}=\overrightarrow{PP''}. \tag{4}$$

如果我们用 $\overrightarrow{PP'''}$ 表示 $\mathbf{B}$，使 $\mathbf{B}$ 和 $\mathbf{A}$ 有同样的起点 P，我们发现 $\mathbf{A}+\mathbf{B}=\overrightarrow{PP''}$ 由顶点为 P,P',P'',P''' (见图 2.1) 的平行四边形的对角线表示 (见图 2.1).

---

1) 有时用记号 $P'-P$ 表示向量 $\overrightarrow{PP'}$. 按照公式 (3b)，这启示着把向量作为点的差这个看法.

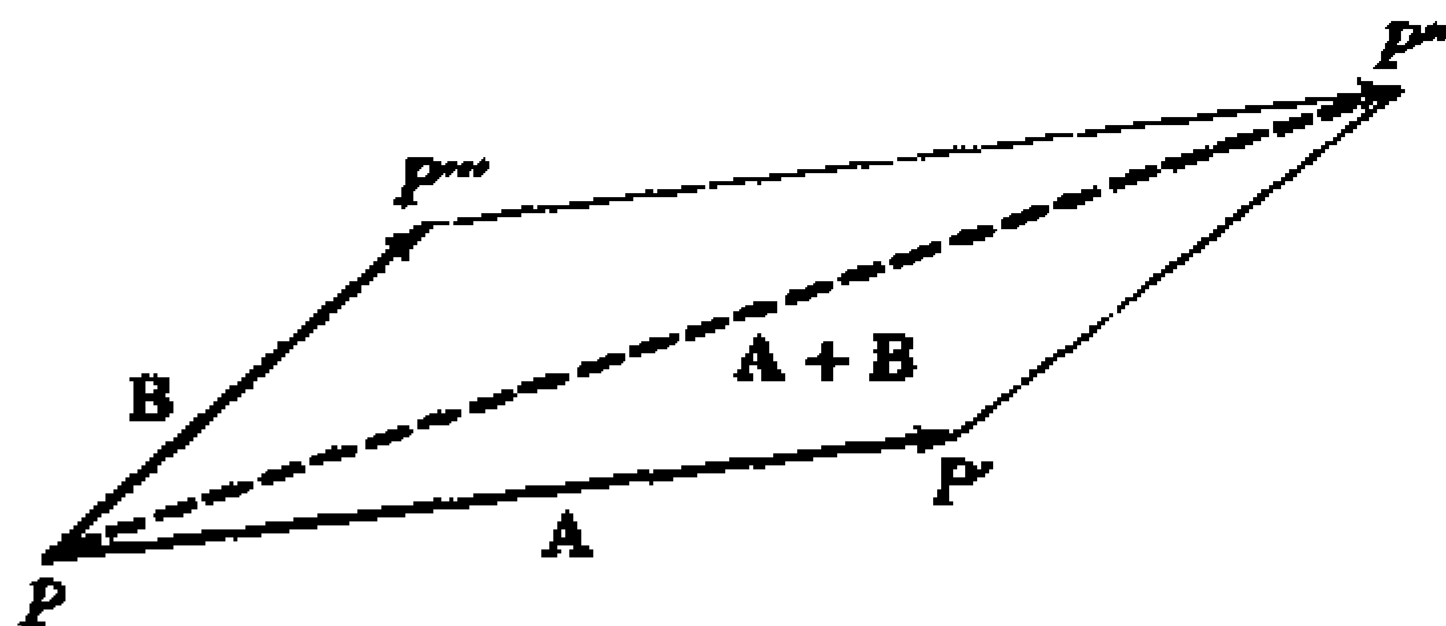


图 2.1 向量的加法

交换向量 $\mathbf{A}=\overrightarrow{PP'}=(x_1'-x_1,x_2'-x_2,\cdots,x_n'-x_n)$ 的起点与终点，得到反向向量

$$\overrightarrow{P'P}=(x_1-x_1',x_2-x_2',\cdots,x_n-x_n')=(-1)\mathbf{A}=-\mathbf{A}.$$

对应于 $-\mathbf{A}$ 的映射 $P'\to P$ 是映射 $\mathbf{A}$ 的逆 (映射;) 先执行 $\mathbf{A}$ 然后再执行 $-\mathbf{A}$, 其结果是恒等映射，它和以下公式是一致的:

$$(-\mathbf{A})+\mathbf{A}=(-1+1)\mathbf{A}=0\mathbf{A}=\mathbf{O}.$$

相应于 (4) 式，对于起点相同的两个向量 $\mathbf{A}=\overrightarrow{PP'}$ 与 $\mathbf{B}=\overrightarrow{PP''}$, 我们可以有常用的向量差的公式

$$\mathbf{B}-\mathbf{A}=\overrightarrow{PP''}-\overrightarrow{PP'}=\overrightarrow{PP''}+\overrightarrow{P'P}=\overrightarrow{P'P}+\overrightarrow{PP''}=\overrightarrow{P'P''}. \quad (4a)$$

这里向量 $\overrightarrow{PP''}$ 与 $\overrightarrow{PP'}$ 的差是由以 P,P',P'' 为顶点的三角形的第三边来表示的.

对于每一个点 $P=(x_1,x_2,\cdots,x_n)$, 我们能联系一个以原点为起点，以 P 点为终点的向量，这就是向量

$$\overrightarrow{OP}=(x_1,x_2,\cdots,x_n),$$

称为 P 点的位置向量. P 点的位置向量的分量正好是 P 点的坐标. 例如，公式 (2a) 中的坐标向量 $\mathbf{E}_i=(0,\cdots,0,1,0,\cdots,0)$ 就是正 x_i 轴上和原点距离为 1 的点的位置向量. 任何一个向量 $\mathbf{A}=\overrightarrow{PP'}$, 总可以写成它的终点与起点的位置向量之差:

$$\overrightarrow{PP'}=\overrightarrow{OP'}-\overrightarrow{OP} \quad (5)$$

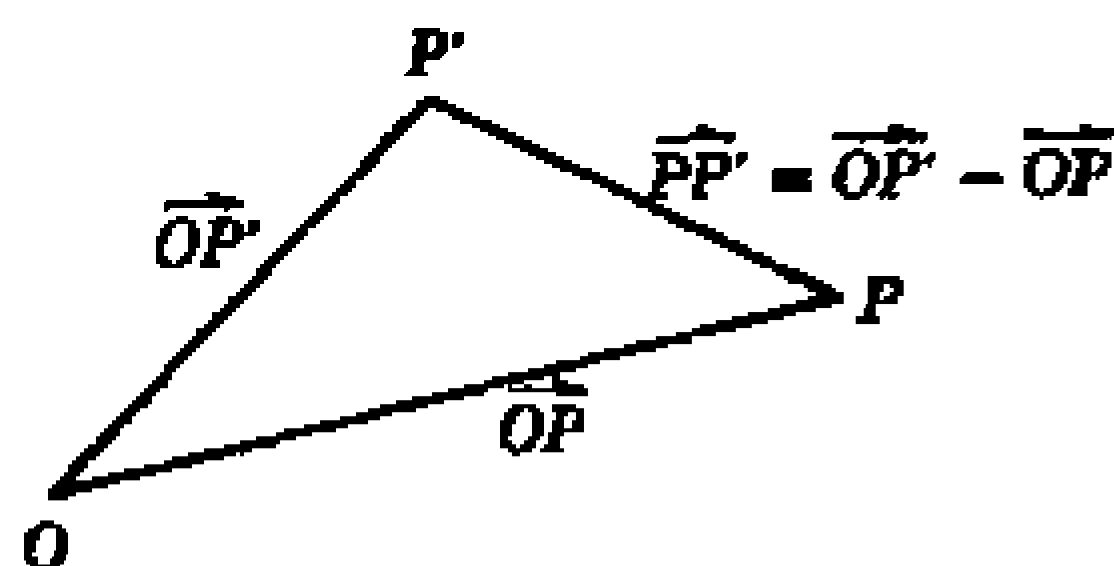


图 2.2 向量 $\overrightarrow{PP'}$ 作为位置向量的差

(见图 2.2).

c. 向量的长度，方向夹角

在 n 维欧氏空间 R^n 中，两个点 $P=(x_1,x_2,\cdots,x_n)$ 和 $P'=(x'_1,x'_2,\cdots,x'_n)$ 的距离由公式 [1)]

$$r=\sqrt{(x'_1-x_1)^2+(x'_2-x_2)^2+\cdots+(x'_n-x_n)^2} \tag{6}$$

给出. 因为在距离 r 的表达式中，仅仅出现 P,P' 的相应坐标之差，我们看出：同一向量 $\mathbf{A}$ 对应的所有的点偶 P,P' 之间的距离都是相同的. 我们称 r 为**向量 A 的长度**，并且记作 $r=|\mathbf{A}|$. 向量 $\mathbf{A}=(a_1,a_2,\cdots,a_n)$ 具有长度

$$|\mathbf{A}|=\sqrt{a_1^2+a_2^2+\cdots+a_n^2}. \tag{6a}$$

零向量 $\mathbf{O}=(0,0,\cdots,0)$ 的长度为 0, 任何其他向量的长度都是正数.

在欧几里得空间中，角度能通过长度来表示. 这可由三角公式(“余弦定律”)得到，这公式给出三角形的三个边 a,b,c 以及 a,b 之间的夹角 γ 之间的关系：

$$\cos\gamma=\frac{a^2+b^2-c^2}{2ab}. \tag{6b}$$

---

1) 在二维或三维的情形，这公式可以应用勾股弦定理从几何上导出. 在高维的情形，r 的表达式可以当作 n 维欧氏空间中，相对于同一笛卡儿坐标系的两个点间的**距离的定义**.

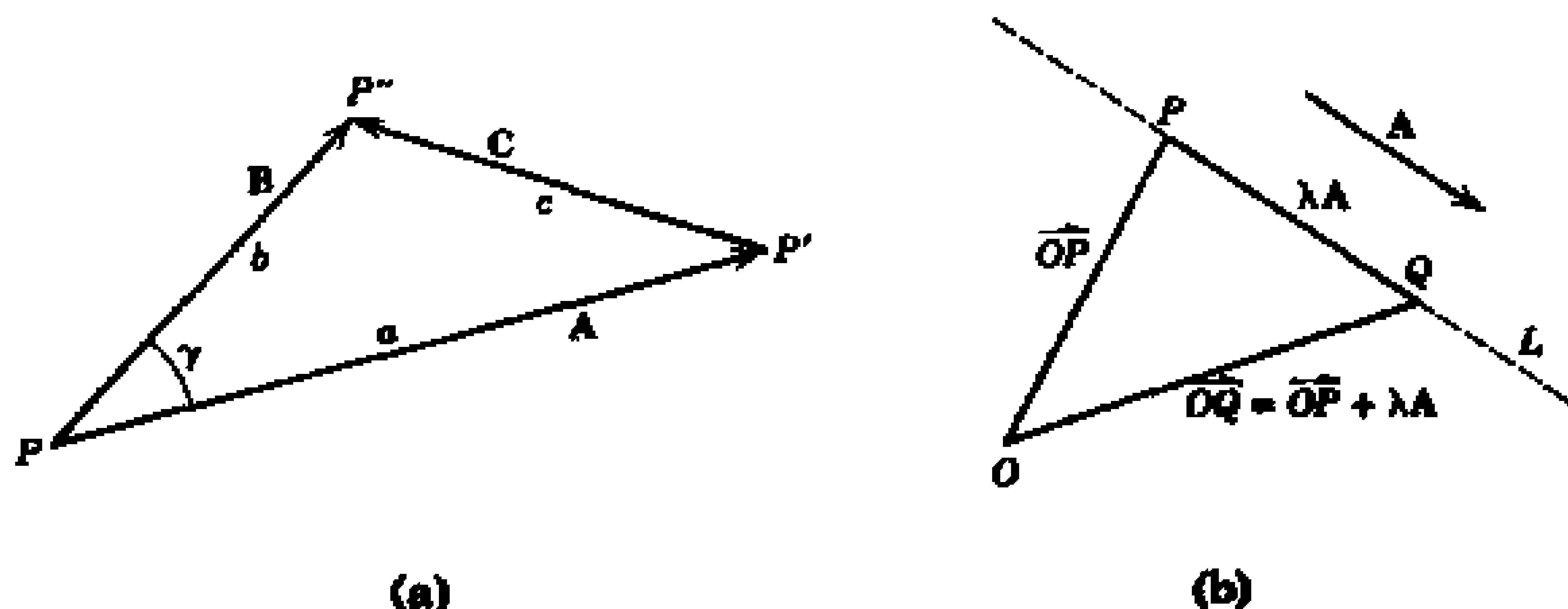


图 2.3 通过一个给定的点具有一个给定的方向的一条直线的向量表示

我们应用这个公式于一个以 P, P', P'' 为顶点的三角形 (图 2.3a). 这三角形的 a 边和 b 边是向量 $\mathbf{A} = \overrightarrow{PP'}$, $\mathbf{B} = \overrightarrow{PP''}$ 的长度, 另一边 c 是向量

$$\mathbf{C} = \overrightarrow{P'P''} = \overrightarrow{PP''} - \overrightarrow{PP'} = \mathbf{B} - \mathbf{A}$$

的长度.

对于

$$\mathbf{A} = (a_1, a_2, \cdots, a_n), \quad \mathbf{B} = (b_1, b_2, \cdots, b_n)$$

我们有

$$\mathbf{C} = (c_1, c_2, \cdots, c_n) = (b_1 - a_1, b_2 - a_2, \cdots, b_n - a_n).$$

由 (6b)

$$\cos\gamma = \frac{|\mathbf{A}|^2 + |\mathbf{B}|^2 - |\mathbf{C}|^2}{2|\mathbf{A}||\mathbf{B}|},$$

其中

$$|\mathbf{A}|^2 = \sum_{i=1}^{n} a_i^2, \quad |\mathbf{B}|^2 = \sum_{i=1}^{n} b_i^2, \quad |\mathbf{C}|^2 = \sum_{i=1}^{n} (b_i - a_i)^2.$$

因此, 对于 $\mathbf{A} \neq 0$, $\mathbf{B} \neq 0$, 有

$$\cos\gamma = \frac{a_1 b_1 + a_2 b_2 + \cdots + a_n b_n}{\sqrt{a_1^2 + \cdots + a_n^2} \cdot \sqrt{b_1^2 + \cdots + b_n^2}}. \tag{7}$$

我们看到三角形 $PP'P''$ 的内角 γ 仅仅依赖于向量 $\mathbf{A}=\overrightarrow{PP'}$ 和 $\mathbf{B}=\overrightarrow{PP''}$. 因此, 我们称由公式 (7) 给出的量 $\cos\gamma$ 为 向量 $\mathbf{A}=(a_1,a_2,\cdots,a_n)$ 与 $\mathbf{B}=(b_1,b_2,\cdots,b_n)$ 之间夹角 [1)]的余弦.

对于 $\cos\gamma$ 的公式 (7), 在任何两个非零向量 $\mathbf{A},\mathbf{B}$ 之间, 实际上总是定义一个 实 角 γ. 因为它永远给出一个值适合 $|\cos\gamma|\leq 1$. 它是 柯西-施瓦茨不等式

$$
\begin{aligned}
&(a_1b_1+a_2b_2+\cdots+a_nb_n)^2\\
\leq&\left(a_1^2+a_2^2+\cdots+a_n^2\right)\cdot\left(b_1^2+b_2^2+\cdots+b_n^2\right)
\end{aligned}
\tag{8}
$$

的一个直接推论.

为从公式 (7) 计算向量 $\mathbf{A}$ 和任意一个向量 $\mathbf{B}$ 之间的夹角, 我们只须知道这些量

$$
\xi_i=\frac{a_i}{\sqrt{a_1^2+a_2^2+\cdots+a_n^2}}\quad(i=1,2,\cdots,n),\tag{9}
$$

它们被称为向量 $\mathbf{A}$ 的 方向余弦. 具有相同方向余弦的所有非零向量都与其他向量作成相同的角度, 因此就说它们具有 相同的方向. 从公式 (7) 看出, $\mathbf{A}$ 的方向余弦可以解释成某个角度的余弦:

$$
\xi_i=\cos\alpha_i,\tag{10}
$$

其中 α_i 是 $\mathbf{A}$ 与第 i 个 "坐标向量"$\mathbf{E}_i=(0,\cdots,0,1,0,\cdots,0)$ 之间的夹角. 向量 $\mathbf{A}$ 的 n 个方向余弦满足以下恒等式 :[2)]

$$
\cos^2\alpha_1+\cos^2\alpha_2+\cdots+\cos^2\alpha_n=1.\tag{11}
$$

---

1) 只当我们限制 γ 在区间 $0\leq\gamma\leq\pi$ 中, 角 γ 才是唯一确定的. 用 $2n\pi\pm\gamma$ 代替 γ (n 是一个整数), 我们将得到所有余弦值相同的角度, 因而它们中的任何一个都可视为向量 $\mathbf{A}$ 与 $\mathbf{B}$ 之间的夹角.

2) 在二维空间中, 关系式 $\cos^2\alpha_1+\cos^2\alpha_2=1$ 允许我们选择 α_2 的值为 $\frac{\pi}{2}-\alpha_1$. 在三维和高维的情形, 方向余弦之间的关系式 (11) 并不对应于角度 α_i 之间的任何简单的线性关系.

唯一没有方向余弦(即没有方向)的向量是零向量.

两个非零向量 $\mathbf{A}$ 与 $\mathbf{B}$ 具有相同的方向，当且仅当它们有相同的方向余弦，也就是

$$\frac{1}{|\mathbf{A}|}\mathbf{A}=\frac{1}{|\mathbf{B}|}\mathbf{B}$$

成立. 显然，这就是当且仅当 $\mathbf{A}$ 与 $\mathbf{B}$ 满足关系 $\mathbf{A}=\lambda\mathbf{B}$, λ 为一正数的情形. 这里 $\lambda=|\mathbf{A}|/|\mathbf{B}|$ 是向量的长度之比. 一个长度为 1 的向量叫做单位向量. 向量

$$(\xi_1,\xi_2,\cdots,\xi_n)=\frac{1}{|\mathbf{A}|}\mathbf{A}$$

的分量为向量 $\mathbf{A}$ 的方向余弦，它是 $\mathbf{A}$ 的方向上的单位向量.

和 $\mathbf{A}$ 相反的向量 $-\mathbf{A}=(-a_1,-a_2,\cdots,-a_n)$ 具有方向余弦 $-\xi_i$. 我们说它的方向与 $\mathbf{A}$ 的方向相反. 当两个非零向量 $\mathbf{A}$ 与 $\mathbf{B}$ 是同方向或反方向时，我们说它们是平行的. 于是平行的一个必要条件是 $\mathbf{A}=\lambda\mathbf{B}$, 其中 λ 是任意一个不等于零的数. 平行于一个给定的方向的任何一个非零向量 $\mathbf{A}$ 的分量 $a_1,a_2,\cdots,a_n$, 叫做该方向的方向数.

如果我们把单位向量 $(\xi_1,\xi_2,\cdots,\xi_n)$ 的起点定在坐标原点 O, 那么终点 $P=(\xi_1,\xi_2,\cdots,\xi_n)$ 就在"单位球面"(即以原点为中心，以 1 为半径的球面) $\xi_1^2+\xi_2^2+\cdots+\xi_n^2=1$ 上. 因为在任何一个给定的方向上恰好有一个单位向量，可见，在 n 维空间中，各方向能由单位球面上的相应点来表示. 球面上相应于相反方向的点位于直径的两端.

一条直线可以直观地想像为一条"常向"曲线，这提示一条 n 维空间的直线可以定义做具有如下性质的点的轨迹，即凡是起点和终点在它上面的向量都是平行的. 这个定义立刻引导出直线的向量表示法. 对于直线 L 上任何不同的两点 P 和 Q, 向量 $\overrightarrow{PQ}$ 都

平行于一个固定的向量 $\mathbf{A}$, 就是说

$$\overrightarrow{PQ} = \lambda \mathbf{A} \quad (\lambda \neq 0).$$

如果我们将 P 点和向量 $\mathbf{A}$ 保持固定而让 Q 跑遍这直线 L 上所有的点，那么对于 Q 点的位置向量我们有下式 (见图 2.3b)

$$\overrightarrow{OQ} = \overrightarrow{OP} + \overrightarrow{PQ} = \overrightarrow{OP} + \lambda \mathbf{A}. \tag{12}$$

这里参量 λ 跑遍所有的实数；数值 $\lambda = 0$ 对应着点 $Q = P$. 如果 Q 点有坐标 $x_1, x_2, \cdots, x_n$; P 点有坐标 $y_1, y_2, \cdots, y_n$; 向量 $\mathbf{A}$ 有分量 $a_1, a_2, \cdots, a_n$, 则公式 (12) 对应着这直线的 **参量表达式**

$$x_i = y_i + \lambda a_i \quad (i = 1, 2, \cdots, n),$$

其中参量 λ 跑遍所有的实数值. 点 P 把直线 L 分成两条半直线，或称为 "射线", 它们是由 λ 的符号来区分的. 对应于 $\lambda > 0$, 向量 $\overrightarrow{PQ}$ 与 $\mathbf{A}$ 方向相同 (指向 $\mathbf{A}$ 的方向); 对应于 $\lambda < 0$ 向量 $\overrightarrow{PQ}$ 与 $\mathbf{A}$ 的方向相反.

d. 向量的数量积

出现在关于两个向量 $\mathbf{A} = (a_1, a_2, \cdots, a_n)$ 与 $\mathbf{B} = (b_1, b_2, \cdots, b_n)$ 之间夹角 γ 的公式 (7) 中分子上的量，叫作向量 $\mathbf{A}$ 与 $\mathbf{B}$ 的数量积，用 $\mathbf{A} \cdot \mathbf{B}$ 表示:

$$\mathbf{A} \cdot \mathbf{B} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n. \tag{13}$$

由几何概念的术语来表达，它能写成

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}||\mathbf{B}| \cos \gamma. \tag{14}$$

两个向量的数量积是它们的长度与它们的方向之间的夹角的余弦的乘积. 如果 $\mathbf{A} = \overrightarrow{PP''}$, $\mathbf{B} = \overrightarrow{PP''}$, 我们能够把 $p = |\mathbf{A}| \cos \gamma$

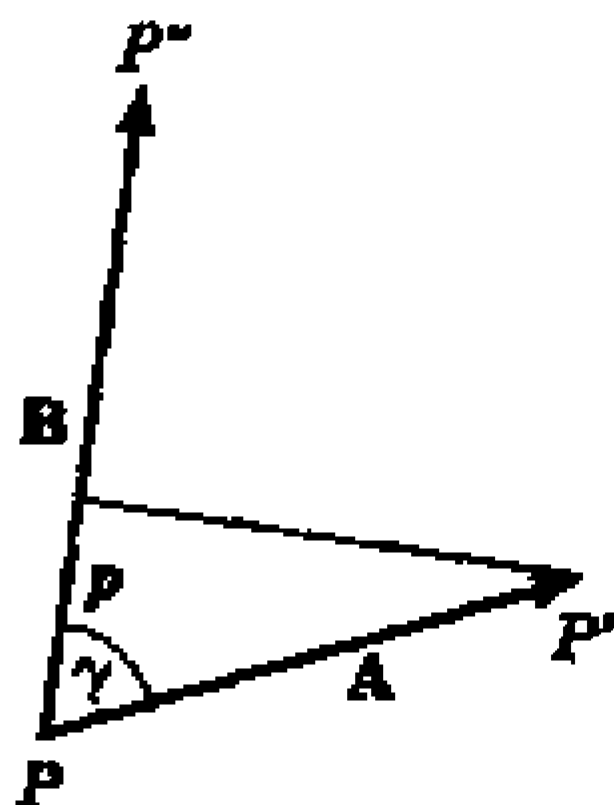


图 2.4 向量 $\mathbf{A}=\overrightarrow{PP'}$ 与 $\mathbf{B}=\overrightarrow{PP''}$ 的数量积

几何地解释成线段 PP' 在直线 PP'' 上的 (带符号的)投影(见图 2.4). 我们称 p 为向量 $\mathbf{A}$ 在向量 $\mathbf{B}$ 方向上的分量. 由公式 (14) 我们有

$$\mathbf{A}\cdot\mathbf{B}=p|\mathbf{B}|. \tag{14a}$$

这样, 向量 $\mathbf{A}$ 与 $\mathbf{B}$ 的数量积等于向量 $\mathbf{B}$ 的长度与向量 $\mathbf{A}$ 在向量 $\mathbf{B}$ 方向上的分量的乘积 [1]. 如果 $\mathbf{B}$ 是正 x_i 轴方向上的坐标向量 $\mathbf{E}=(0,\cdots,1,\cdots,0)$, 则 $\mathbf{A}$ 在 $\mathbf{B}$ 方向上的分量是 a_i, 即 $\mathbf{A}$ 的第 i 个分量.

由定义 (13), 容易验证数量积满足通常的代数定律:

$$\mathbf{A}\cdot\mathbf{B}=\mathbf{B}\cdot\mathbf{A}, \qquad \text{(交换律)} \tag{15a}$$

$$\lambda(\mathbf{A}\cdot\mathbf{B})=(\lambda\mathbf{A})\cdot\mathbf{B}=\mathbf{A}\cdot(\lambda\mathbf{B}), \qquad \text{(结合律)}^{2)} \tag{15b}$$

$$\left.\begin{aligned}\mathbf{A}\cdot(\mathbf{B}+\mathbf{C})&=\mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{C},\\(\mathbf{A}+\mathbf{B})\cdot\mathbf{C}&=\mathbf{A}\cdot\mathbf{C}+\mathbf{B}\cdot\mathbf{C}.\end{aligned}\right\} \quad \text{(分配律)} \tag{15c}$$

数量积的基本重要性来源于下述事实, 它通过向量 $\mathbf{A}$ 与 $\mathbf{B}$ 的分量的表达式, 具有简单的代数表达式 (13), 而与此同时, 它又具有由公式 (14) 所表明的那种纯粹的几何解释 (公式 (14) 不涉及向

1) 显然, 它也等于向量 $\mathbf{A}$ 的长度乘以向量 $\mathbf{B}$ 在 $\mathbf{A}$ 方向上的分量.

2) 因为二向量的数量积不是向量, 而是数, 因此没有包含数量积与第三向量的结合律.

量在任何特定坐标系中的分量). 数量积不仅在描述角度时有用, 它也是推导面积和体积的分析表达式的基础.

从柯西 - 施瓦茨不等式 (8), 我们推出数量积满足不等式

$$|\mathbf{A}\cdot\mathbf{B}| \le |\mathbf{A}||\mathbf{B}|, \tag{16}$$

它正好表明 $|\cos\gamma| \le 1$. 我们将要看到 (第 197 页), (16) 式中的等号仅仅当 $\mathbf{A}$ 与 $\mathbf{B}$ 互相平行或者它们之中至少有一个是零向量的时候才成立.

我们注意公式 (13), 当 $\mathbf{B} = \mathbf{A}$ 时, 按公式 (6a) 有

$$\mathbf{A}\cdot\mathbf{A} = |\mathbf{A}|^2, \tag{17a}$$

这就是说, 一个向量与它自己的数量积是它的长度的平方. 这也可以从公式 (14) 得到, 因为向量与它自己的夹角为零, 即 $\gamma = 0$. 对于非零向量的重要关系式

$$\mathbf{A}\cdot\mathbf{B} = 0 \tag{17b}$$

对应着 $\cos\gamma = 0$ 或 $\gamma = \dfrac{\pi}{2}$. 因此, 关系式 (17b) 刻划了向量 $\mathbf{A}$ 与 $\mathbf{B}$ 互相 "垂直" 或互相 "正交" 这个特征. 另一方面, $\mathbf{A}\cdot\mathbf{B} > 0$ 意味着 $\cos\gamma > 0$; 就是说, 我们能够确定一个值 γ 满足 $0 \le \gamma < \dfrac{\pi}{2}$; 这两个向量的方向形成一个锐角. 类似地, $\mathbf{A}\cdot\mathbf{B} < 0$ 意味着这两个向量相互形成一个钝角, $\dfrac{\pi}{2} < \gamma \le \pi$.

例如, 两个坐标向量 (见第 133 页)

$$\mathbf{E}_1 = (1, 0, \cdots, 0) \text{ 与 } \mathbf{E}_2 = (0, 1, 0, \cdots, 0)$$

是互相垂直的, 因为 $\mathbf{E}_1\cdot\mathbf{E}_2 = 1\cdot 0 + 0\cdot 1 + 0\cdot 0 + \cdots + 0\cdot 0 = 0$. 更一般地, 任何两个不同的坐标向量 $\mathbf{E}_i$ 与 $\mathbf{E}_k$ 都是垂直的:

$$\mathbf{E}_i\cdot\mathbf{E}_k = 0, \quad (i \ne k). \tag{17c}$$

对于 $k=i$, 我们显然有

$$\mathbf{E}_i\cdot\mathbf{E}_i=|\mathbf{E}_i|^2=1;$$

即坐标向量的长度是1.

e. 超平面方程的向量形式

在 n 维空间 R^n 中，满足形如

$$a_1x_1+a_2x_2+\cdots+a_nx_n=c \tag{18}$$

的线性方程 (其中 $a_1,a_2,\cdots,a_n$ 不全为零) 的点的轨迹叫作一个超平面. 字头 "超" 是需要的，因为 n 维空间包含着各种维的 "平面" 或 "线性流型"; 而超平面可等同于包含在 n 维空间 R^n 中的 $n-1$ 维欧氏空间. 它们是通常三维空间中的二维平面，平面内的直线，直线上的点.

引进向量 $\mathbf{A}=(a_1,a_2,\cdots,a_n)$ 与 P 点的位置向量 $\mathbf{X}=(x_1,x_2,\cdots,x_n)=\overrightarrow{OP}$, 我们能将方程 (18) 写成向量形式

$$\mathbf{A}\cdot\mathbf{X}=c\quad(\mathbf{A}\neq 0). \tag{18a}$$

设 $\mathbf{Y}=(y_1,y_2,\cdots,y_n)=\overrightarrow{OQ}$ 是超平面上一个特定点 Q 的位置向量，使得 $\mathbf{A}\cdot\mathbf{Y}=c$. 从 (18a) 减去这个等式，就发现超平面上的点 P 满足

$$0=\mathbf{A}\cdot\mathbf{X}-\mathbf{A}\cdot\mathbf{Y}=\mathbf{A}\cdot(\mathbf{X}-\mathbf{Y})=\mathbf{A}\cdot\overrightarrow{PQ}. \tag{19}$$

因此向量 $\mathbf{A}$ 垂直于超平面上任意两点的连线. 从超平面上任何一点 Q 出发，沿着垂直于 $\mathbf{A}$ 的所有方向行进，所得的点就组成这个超平面. 我们称 $\mathbf{A}$ 的方向为超平面的法方向 (见图 2.5).

具有方程 (18a) 的超平面，把空间分成两个开的半空间，由 $\mathbf{A}\cdot\mathbf{X}<c$ 和 $\mathbf{A}\cdot\mathbf{X}>c$ 给出. 向量 $\mathbf{A}$ 指向半空间 $\mathbf{A}\cdot\mathbf{X}>c$ 的一边.

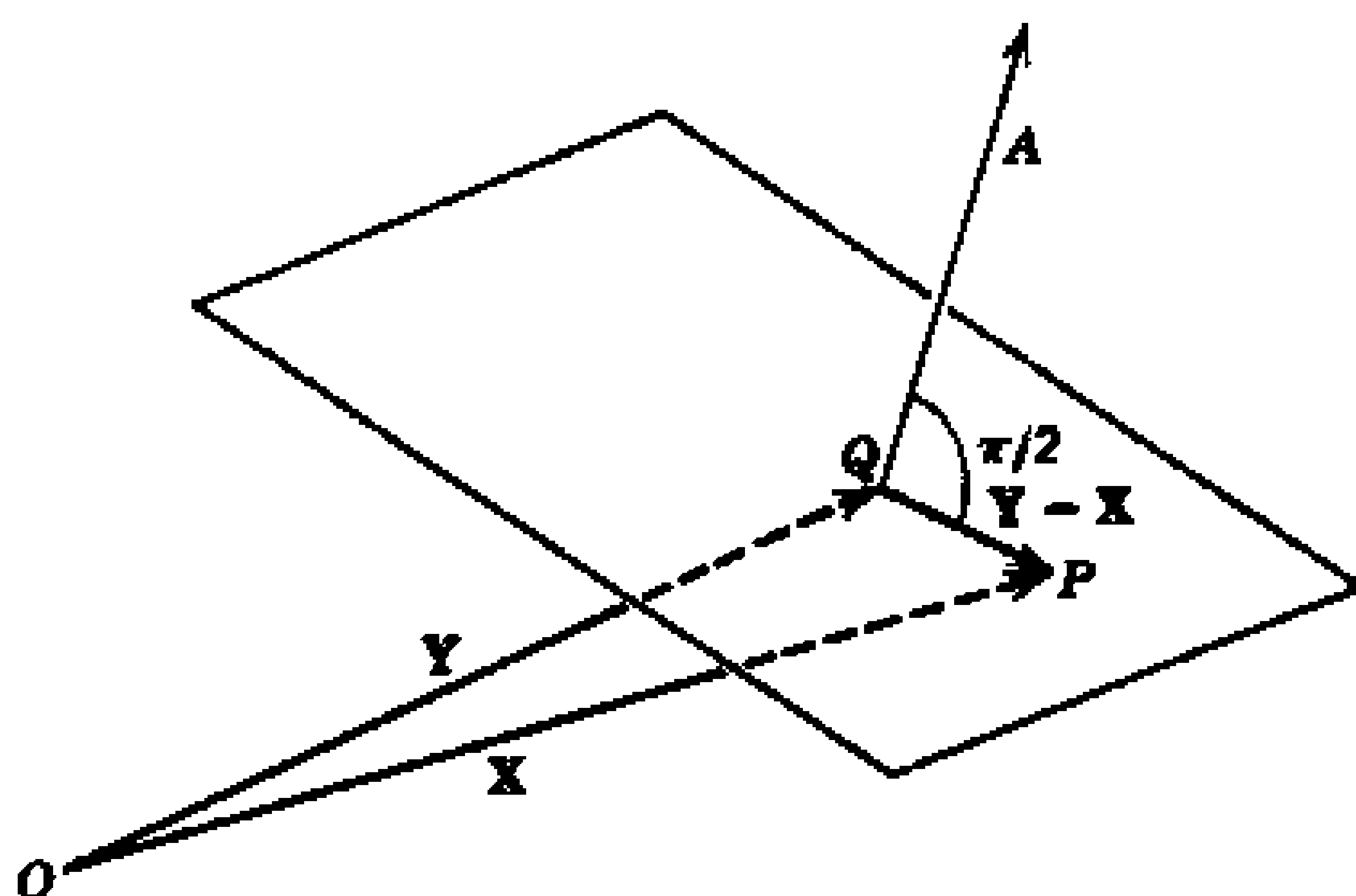


图 2.5　超平面的形成

这指的是，超平面上以一点 Q 出发的在 $\mathbf{A}$ 方向上的射线是由这样一类点组成，它们的位置向量 $\mathbf{X}$ 满足 $\mathbf{A}\cdot\mathbf{X}>c$. 事实上，这样的射线上的点 P 的位置向量 $\mathbf{X}$ 由

$$\mathbf{X}=\overrightarrow{OP}=\overrightarrow{OQ}+\lambda\mathbf{A}=\mathbf{Y}+\lambda\mathbf{A}$$

给出 (见 (12)), 其中 $\mathbf{Y}$ 是 Q 点的位置向量，λ 为一正数. 所以显然有

$$\mathbf{A}\cdot\mathbf{X}=\mathbf{A}\cdot\mathbf{Y}+\mathbf{A}\cdot\lambda\mathbf{A}=c+\lambda|\mathbf{A}|^2>c.$$

更一般地，任何一个同 $\mathbf{A}$ 作成锐角的向量 $\mathbf{B}$ 都指向半空间 $\mathbf{A}\cdot\mathbf{X}>c$, 因为由 $\mathbf{A}\cdot\mathbf{B}>0$ 可推出

$$\mathbf{A}\cdot\mathbf{X}=\mathbf{A}\cdot(\mathbf{Y}+\lambda\mathbf{B})=\mathbf{A}\cdot\mathbf{Y}+\lambda\mathbf{A}\cdot\mathbf{B}>c.$$

如果常数 c 是正的，因为 $\mathbf{A}\cdot\mathbf{O}=0<c$, 则半空间 $\mathbf{A}\cdot\mathbf{X}<c$ 将要包含原点，这时 $\mathbf{A}$ 具有背离原点的方向.

描述一个给定的超平面的线性方程 (18a) 不是唯一的. 因为我们能够将方程乘上一个任意常数因子 $\lambda\neq 0$, 那相当于用平行于向量 $\mathbf{A}$ 的向量 $\lambda\mathbf{A}$ 代替 $\mathbf{A}$, 用 λc 代替 c. 如果 $c\neq 0$, 也就是说，如果超平面不经过原点，我们就能选取

$$\lambda=\frac{\operatorname{sgn} c}{|\mathbf{A}|}.$$

用 λ 乘 (18a), 我们得到超平面的法式方程

$$\mathbf{B}\cdot\mathbf{X}=p \tag{20}$$

这里 p 是一个正常数, $\mathbf{B}$ 是具有背离原点方向的单位法向量. 方程 (20) 中的常数 p 就是原点到超平面的距离, 也就是, 从原点到超平面上任意点的最短距离. 因为若 P 点是超平面上任意一点, 而 $\mathbf{X}$ 是 P 点的位置向量, 则从原点到 P 点的距离由

$$|\overrightarrow{OP}|=|\mathbf{X}|=|\mathbf{X}|\cdot|\mathbf{B}|$$

给出, 而由 (16) 和 (20) 便推出

$$|\overrightarrow{OP}|\geq\mathbf{B}\cdot\mathbf{X}=p.$$

等号对于超平面上特殊点 P 成立, 其位置向量为

$$\overrightarrow{OP}=\mathbf{X}=p\mathbf{B}.$$

这个点与原点的连线的方向便是超平面的法方向. 更一般地, 我们能够找到由超平面到空间中任何一个具有位置向量 $\mathbf{Y}$ 的点 Q 的距离 d. 读者可以自行验证

$$d=|\mathbf{B}\cdot\mathbf{Y}-p|. \tag{20a}$$

f. 向量的线性相关与线性方程组

数学分析中的很多问题都能化为 n 维空间中的若干向量间的线性关系的研究. 一个向量 $\mathbf{Y}$, 如果能够表示成向量 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_m$ 的线性组合, 就是说, 如果存在着 m 个数 $x_1,x_2,\cdots,x_m$, 使得

$$\mathbf{Y}=x_1\mathbf{A}_1+x_2\mathbf{A}_2+\cdots+x_m\mathbf{A}_m \tag{21}$$

成立, 则称向量 $\mathbf{Y}$ 关于向量 $\mathbf{A}_1,\cdots,\mathbf{A}_m$ 是相关的[1]. 这里 m 是任何一个自然数. 零向量永远是相关的, 因为选择所有的数量 x_i

1) 这里我们称 "相关的", 在文献中常常被称为 "线性相关的". 因为在向量之间, 我们不考虑其他类型的相关, 我们去掉了 "线性" 这个词.

为零时，它能被表示成 (21) 的形式. 向量 $\mathbf{Y}$ 对于一个单个的向量 $\mathbf{A}_1 \neq 0$ 是相关的，意味着向量 $\mathbf{Y} = 0$ 或者 $\mathbf{Y}$ 与 $\mathbf{A}_1$ 平行. 将 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 选成 n 个坐标向量

$$\begin{aligned}&\mathbf{E}_1 = (1,0,\cdots,0), \quad \mathbf{E}_2 = (0,1,0,\cdots,0), \cdots,\\&\mathbf{E}_n = (0,\cdots,0,1),\end{aligned} \tag{22}$$

我们就看到关系式 (21) 对于任何一个向量 $\mathbf{Y} = (y_1, y_2, \cdots, y_n)$ 都成立；只要我们选取 $x_1 = y_1, x_2 = y_2, \cdots, x_n = y_n$，就有

$$\mathbf{Y} = y_1\mathbf{E}_1 + y_2\mathbf{E}_2 + \cdots + y_n\mathbf{E}_n. \tag{23}$$

因此，空间中每一个向量关于坐标向量都是相关的. 另一方面，容易看出 n 个坐标向量 $\mathbf{E}_i$ 中没有一个跟另外一些个相关. 更一般地，如果一个向量 $\mathbf{Y} \neq 0$，但它垂直于向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 中的每一个，则向量 $\mathbf{Y}$ 关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 不能是相关的. 因为，关系式 (21) 与它自己的数量乘积产生

$$\begin{aligned}|\mathbf{Y}|^2 = \mathbf{Y}\cdot\mathbf{Y} &= \mathbf{Y}\cdot(x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_m\mathbf{A}_m)\\&= x_1\mathbf{Y}\cdot\mathbf{A}_1 + x_2\mathbf{Y}\cdot\mathbf{A}_2 + \cdots + x_m\mathbf{Y}\cdot\mathbf{A}_m = 0,\end{aligned}$$

所以有 $\mathbf{Y} = \mathbf{O}$.

我们称向量组 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 是相关的，如果存在不全为零的数 $x_1, x_2, \cdots, x_m$，使得

$$x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_m\mathbf{A}_m = \mathbf{O} \tag{24}$$

成立. 如果 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 不是相关的 —— 也即，(24) 只对 $x_1 = x_2 = \cdots = x_m = 0$ 成立 —— 我们就称 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 是无关的. 例如，坐标向量 $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n$ 是无关的，因为

$$\mathbf{O} = x_1\mathbf{E}_1 + x_2\mathbf{E}_2 + \cdots + x_n\mathbf{E}_n = (x_1, x_2, \cdots, x_n)$$

显然隐含着 $x_1 = x_2 = \cdots = x_n = 0$.

"一个向量关于一组向量相关"和"一组向量相关"这两个概念是紧密联系着的. 几个向量是相关的，当且仅当它们中的一个关于其余的是相关的. 因为，表明 $\mathbf{Y}$ 关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 相关的关系式 (21) 显然能写成如下形式:

$$x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_m\mathbf{A}_m + (-1)\mathbf{Y} = 0,$$

这表明 $m+1$ 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$, $\mathbf{Y}$ 是相关的. 反之，如果 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 是相关的，我们就有一个形如 (24) 的关系式，其中系数 x_i 不全为零. 若 $x_k \neq 0$, 我们就能够对 $\mathbf{A}_k$ 解方程 (24), 而将 $\mathbf{A}_k$ 表示成其余向量的线性组合.

向量 $\mathbf{Y}$ 关于向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的相关性意味着，一个特定的线性方程组有解 $x_1, x_2, \cdots, x_m$. 因为让 $\mathbf{Y} = (y_1, y_2, \cdots, y_n)$, 而让向量 $\mathbf{A}_k$ 写成分量如

$$\mathbf{A}_k = (a_{1k}, a_{2k}, \cdots, a_{nk}),$$

则用分量写出的向量方程 (21), 等价于关于未知量 $x_1, x_2, \cdots, x_m$ 的 n 个线性方程组成的方程组

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= y_2, \\ \cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= y_n. \end{aligned} \tag{25}$$

显然，若要 $\mathbf{Y}$ 关于 $\mathbf{A}, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 是相关的，必须且只须方程组 (25) 至少有一组解 $x_1, x_2, \cdots, x_m$. 类似地，向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 是相关的，当且仅当齐次方程组

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= 0, \end{aligned} \tag{25a}$$

$$\cdots\cdots$$

$$a_{n1}x_1+a_{n2}x_2+\cdots+a_{nm}x_m=0$$

有一组"非平凡"解 $x_1,x_2,\cdots,x_m$，即，有一组解不同于平凡解 [1)]

$$x_1=x_2=\cdots=x_m=0.$$

在 n 维空间中，我们已经找到了一组无关的 n 个向量，就是坐标向量 $\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_n$. 向量理论的基础是这个事实，即 n 是无关向量的最大个数.

线性相关基本定理　在 n 维空间中，任意 $n+1$ 个向量都是相关的.

在证明这个定理之前，我们先考虑它的一些深刻的含义. 我们能立刻由它推出，在 n 维空间中，任何多于 n 个的一组向量都是相关的. 因为任何 m 个向量中的前 $n+1$ 个向量的相关式 (24)，能够被看作是 m 个向量的相关式，只要我们假定剩下的向量的系数都是零. 于是，基本定理隐含着：在齐次线性方程组(25a)中，如果 $m>n$，即未知数的个数超过方程式的个数时，它永远有非平凡解.

我们能够用不同的方式几何地表达最后这个陈述，如果我们把方程组 (25a) 的每一个方程解释为 m 维空间的某两个向量的数量积为零. 于是一组非平凡解 $x_1,x_2,\cdots,x_m$ 就对应到一个向量 $\mathbf{X}=(x_1,x_2,\cdots,x_m)\neq 0$. 两个非零向量的数量积为零意味着这两个向量是互相垂直的. 方程组 (25a) 表明 $\mathbf{X}$ 垂直于这 n 个向量 $(a_{11},a_{12},\cdots,a_{1m}),(a_{21},a_{22},\cdots,a_{2m}),\cdots,(a_{n1},a_{n2},\cdots,a_{nm})$. 于是我们有：给定一组非零向量，它们的个数少于空间的维数，则我们能够找到一个向量与它们中所有的向量都垂直(由第 147 页知道，该向量与它们是不相关的).

---

1) 形如 $P(x_1,x_2,\cdots,x_m)=0$ 的方程，其中 P 是齐次多项式 (见第 14 页)，叫齐次方程. 它们永远有平凡解 $x_1=x_2=\cdots=x_m=0$. 另外，任何一组解 $x_1,x_2,\cdots,x_m$，当把所有的 x_i 乘上同一因子 λ 时，它仍旧是一组解.

回到 n 维空间中的向量，我们看到基本定理的一个进一步的推论：**n 维空间中给定 n 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$，只要它们是无关的，则该空间中任何一个向量 $\mathbf{Y}$ 都与它们是相关的**．因为 $n+1$ 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n, \mathbf{Y}$ 必定是相关的，我们有形如

$$z_1\mathbf{A}_1 + z_2\mathbf{A}_2 + \cdots + z_n\mathbf{A}_n + z_{n+1}\mathbf{Y} = \mathbf{0}$$

的关系式，其中 $z_1, z_2, \cdots, z_{n+1}$ 不全为零．于是有 $z_{n+1} \neq 0$，因为否则 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 就会是相关的，这与假设矛盾．由此推出

$$\mathbf{Y} = x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n, \tag{26}$$

式中

$$x_i = -\frac{z_i}{z_{n+1}} \quad (i = 1, 2, \cdots, n).$$

顺便指出，把 $\mathbf{Y}$ 作为无关向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的线性组合的表达式 (26) 中的系数是唯一的，因为，假如存在着第二种表达式

$$\mathbf{Y} = y_1\mathbf{A}_1 + y_2\mathbf{A}_2 + \cdots + y_n\mathbf{A}_n,$$

通过相减就会得到

$$(x_1 - y_1)\mathbf{A}_1 + (x_2 - y_2)\mathbf{A}_2 + \cdots + (x_n - y_n)\mathbf{A}_n = \mathbf{0}.$$

由此我们从向量 $\mathbf{A}_1, \cdots, \mathbf{A}_n$ 的无关性推出所有的系数都为零，因而推出 $x_1 = y_1, \cdots, x_n = y_n$.

另一方面，如果 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是相关的，我们必定能找到一个向量 $\mathbf{Y}$，它与 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是不相关的，因为在这种情形下，向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 中有一个对于其余的是相关的，比如，$\mathbf{A}_n$ 相关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{n-1}$；则相关于向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的一个向量 $\mathbf{Y}$ 关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{n-1}$ 也是相关的．但是，n 维空间中存在向量 $\mathbf{Y}$ 不相关于给定的 $n-1$ 个向量 (见第 149 页).

因为向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的无关性等价于对应的齐次线性方程组 (25a) 只有零解，于是我们可从基本定理推导出线性方程组可解性的基础定理：

线性方程组

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2, \\
&\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= y_n,
\end{aligned} \tag{27}
$$

对于任意给定的 n 个数 $y_1, y_2, \cdots, y_n$ 都有一组唯一的解 $x_1, x_2, \cdots, x_n$ 只需要齐次方程组

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0, \\
&\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0,
\end{aligned} \tag{27a}
$$

仅仅有零解 $x_1 = x_2 = \cdots = x_n = 0$. 如果方程组(27a)有一组非零解，我们就能找到一组值 $y_1, y_2, \cdots, y_n$，对于它们方程组(27)无解.

这是一个纯粹的存在性定理，它并未指出，如果解存在的话，怎样真正得到解 $x_1, x_2, \cdots, x_n$. 这可借助于下面 2.3 节中讨论的行列式来得到.

我们来证明基本定理，对维数 n 用归纳法. 定理说的是，n 维空间中任何 $n+1$ 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n, \mathbf{Y}$ 是相关的. 对于 $n=1$，向量变成数量，要证明的断语变成：对于任何两个数 y 和 A，我们能找到不全为零的数 x_0, x_1，使得

$$x_0 y + x_1 A = 0.$$

这是容易办到的. 如果 $y = A = 0$，我们取 $x_0 = x_1 = 1$；在其他情形，我们取 $x_0 = A$, $x_1 = -y$.

假定我们已经证明了 $n-1$ 维空间中任何 n 个向量是相关的. 令 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$, y 是 n 维空间的向量. 我们要证明 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$, y 是相关的. 如果 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 自身已经是相关的, 则结论肯定成立. 因此, 我们限于考虑 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是不相关的情形. 我们将要证明 $\mathbf{Y}$ 关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是相关的. 这只需证明 (22) 中每一个坐标向量 $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n$ 关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是相关的. 因为, 由 (23) 式, 任何向量 $\mathbf{Y}$ 是 $\mathbf{E}_i$ 的一个线性组合, 因而, 如果 $\mathbf{E}_i$ 能用 $\mathbf{A}_k$ 线性表示, 则 $\mathbf{Y}$ 也可以用 $\mathbf{A}_k$ 线性表示. 我们将只证明 $\mathbf{E}_n$ 关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是相关的, 因为对于其他的 $\mathbf{E}_i$ 的证明是类似的. 我们只要证明线性方程组

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0, \\ \qquad (i = 1, 2, \cdots, n-1) \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 1 \end{cases} \tag{28}$$

有一组解 $x_1, x_2, \cdots, x_n$. 这里前 $n-1$ 个方程是齐次方程, 作为 $n-1$ 维空间中的 n 个向量是相关的这个归纳法假设的推论, 它有一组非零解 $x_1, x_2, \cdots, x_n$. 对于这组解, 令

$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = c.$$

这里 $c \neq 0$, 因为否则 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 就会是相关的. 用 c 除 $x_1, x_2, \cdots, x_n$, 我们就得到方程组 (28) 的解. 这就完成了基本定理的证明.

练 习 2.1

1. 给出通过点 $P = (-2, 0, 4)$ 并且沿着向量 $\mathbf{A} = (2, 1, 3)$ 的方向的直线的坐标表达式.

2. (a) 什么是通过点 $P = (3, -2, 2)$ 和 $Q = (6, -5, 4)$ 的直线方程?

(b) 给出通过任意两个不同的点 P 和 Q 的直线方程.

3. 如果 $\mathbf{A}$ 和 $\mathbf{B}$ 是起点为 O, 终点分别为 P, Q 的两个向量, 则以 O 为起点, 以分 PQ 为比值 $\lambda : (1-\lambda)$ 的点为终点的向量被

线性组合

$$(1-\lambda)\mathbf{A}+\lambda\mathbf{B}$$

给出.

4. 在练习第 3 题中，当 λ 取何值时，其位置向量所对应的点位于从 P 到 Q 的射线上？

5. 四面体 $PQRS$ 顶点的质心可以定义为分 MS 为 1 : 3 的点，此处 M 是 PQR 顶点的质心. 证明这个定义与顶点的次序无关，并且与质心的通常的定义是一致的.

6. 四面体的两个没有公共顶点的边叫作对边. 例如习题 5 中的四面体的边 PQ 和 RS 是对边. 证明四面体对边中点的连线通过顶点的质心.

7. 设 $A_1, A_2, \cdots, A_n$ 是空间的任意 n 个质点，分别具有质量 $m_1, m_2, \cdots, m_n$; G 是它们的质心. 用 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 表示起点为 G 终点为 $A_1, A_2, \cdots, A_n$ 的向量. 证明

$$m_1\mathbf{A}_1+m_2\mathbf{A}_2+\cdots+m_n\mathbf{A}_n=0.$$

8. 实数形成一个一维向量空间，其中“向量”的加法是通常的加法，“向量”与数量的乘法是通常的乘法. 证明正实数也形成一个向量空间，其中向量的加法是通常的乘法，而数量的乘法则适当地给予定义.

9. 证明复数作成二维向量空间，其中加法是通常的加法，而数量则是实数.

10. 设 P 和 Q 是球面上对立于直径两端的点，R 为球面上其他任意一点. 证明 PR 与 QR 成直角.

11. (a) 求经过 $P=(-3,2,1)$ 垂直于向量 $\mathbf{A}=(1,2,-2)$ 的平面的法式方程.

(b) 从这个平面到点 $Q=(1,-1,-1)$ 的距离是什么？

(c) 问 O 与 Q 在平面的同侧还是异侧？

12. (a) 设超平面方程由 (18) 给出. 确定从一点 P 到超平面的垂足的坐标.

(b) 在习题 11 中，给出 O 和 Q 到平面垂足的坐标.

13. 设 $\mathbf{A}$ 和 $\mathbf{B}$ 是不平行的向量. 证明向量

$$\mathbf{C}=\mathbf{A}-\frac{\mathbf{A}\cdot\mathbf{B}}{|\mathbf{B}|^2}\mathbf{B}$$

与 $\mathbf{B}$ 垂直. 向量 $\mathbf{C}$ 称为 $\mathbf{A}$ 的垂直于 $\mathbf{B}$ 的分量.

14. 找到平面

$$\mathbf{A}x+\mathbf{B}y+\mathbf{C}z+\mathbf{D}=0$$

与直线

$$x=x_0+\alpha t,\quad y=y_0+\beta t,\quad z=z_0+\gamma t$$

之间的夹角.

2.2 矩阵与线性变换

a. 基的变换，线性空间

在 n 维空间 R^n 中，每一个向量 $\mathbf{Y}$ 能够写成 (22) 中定义的坐标向量 $\mathbf{E}_1,\cdots,\mathbf{E}_n$ 的线性组合，即

$$\mathbf{Y}=y_1\mathbf{E}_1+y_2\mathbf{E}_2+\cdots+y_n\mathbf{E}_n,\tag{29}$$

其中 y_i 是向量 $\mathbf{Y}$ 的分量. 我们能把坐标向量与分量的概念推广到 S_n 中任意 m 个无关的向量 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_m$. 如果 $\mathbf{Y}$ 是一个与 $\mathbf{A}_i$ 相关的向量，我们有

$$\mathbf{Y}=x_1\mathbf{A}_1+x_2\mathbf{A}_2+\cdots+x_m\mathbf{A}_m,\tag{30}$$

其中系数 x_i 由 $\mathbf{Y}$ 唯一确定. 我们称 $x_1,x_2,\cdots,x_m$ 为 $\mathbf{Y}$ 关于基 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_m$ 的分量. 对于这组基，基向量 $\mathbf{A}_1$ 有分量 $1,0,\cdots,0$; 基向量 $\mathbf{A}_2$ 有分量 $0,1,0,\cdots,0$; 等等. 对于任意数量 λ, 向量

$$\lambda\mathbf{Y}=\lambda x_1\mathbf{A}_1+\lambda x_2\mathbf{A}_2+\cdots+\lambda x_m\mathbf{A}_m$$

也是与 $\mathbf{A}_i$ 相关的，并且有分量 $\lambda x_1, \lambda x_2, \cdots, \lambda x_m$. 类似地，如果

$$\mathbf{Y}' = x_1'\mathbf{A}_1 + x_2'\mathbf{A}_2 + \cdots + x_m'\mathbf{A}_m$$

是相关于 $\mathbf{A}_i$ 的第二个向量，则二向量的和

$$\mathbf{Y} + \mathbf{Y}' = (x_1 + x_1')\mathbf{A}_1 + (x_2 + x_2')\mathbf{A}_2 + \cdots + (x_m + x_m')\mathbf{A}_m$$

关于我们的基有分量 $x_1 + x_1', x_2 + x_2', \cdots, x_m + x_m'$.

对于 $m < n$ 并不是 n 维空间中所有的向量都相关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$. 相关于 m 个无关向量的全部向量被称为形成一个 **m维向量空间**. 我们能具体作成这样一个空间，即选取任意一点 P_0 (具有位置向量 $\mathbf{B} = \overrightarrow{OP_0}$), 而以 P_0 为起点作所有的向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$. 令

$$\mathbf{A}_i = \overrightarrow{P_0P_i} \quad (i = 1, 2, \cdots, m), \tag{31a}$$

并且设 $\mathbf{Y} = \overrightarrow{P_0P}$ 是由 (30) 给定的向量. 于是，点 P 具有位置向量

$$\overrightarrow{OP} = \overrightarrow{OP_0} + \overrightarrow{P_0P} = \mathbf{B} + x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_m\mathbf{A}_m. \tag{31b}$$

我们称满足关系式 (31b) 的全部点 P 形成**经过** P_0 **由** $A_1, A_2, \cdots, A_m$ **张成的** m **维线性流形** S_m. S_m 中的每一个点 P 唯一地决定一组值 $x_1, x_2, \cdots, x_m$, 称为点 P 的 **仿射坐标**. 在 S_m 的这个仿射坐标系中，"原点"，即有 $x_1 = x_2 = \cdots = x_m = 0$ 的点，是点 P_0; 仿射坐标 $x_1 = 1, x_2 = \cdots = x_m = 0$ 的点是点 P_1, 即向量 $A_1 = \overrightarrow{P_0P_1}$ 的终点，等等. 并且对于 S_m 中具有位置向量

$$\overrightarrow{OP} = \mathbf{B} + x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_m\mathbf{A}_m,$$
$$\overrightarrow{OP'} = \mathbf{B} + x_1'\mathbf{A}_1 + x_2'\mathbf{A}_2 + \cdots + x_m'\mathbf{A}_m$$

的两个点 P 和 P', 向量

$$\overrightarrow{PP'} = \overrightarrow{OP'} - \overrightarrow{OP}$$

$$= (x'_1 - x_1)\mathbf{A}_1 + (x'_2 - x_2)\mathbf{A}_2 + \cdots + (x'_m - x_m)\mathbf{A}_m$$

关于基 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的分量是点 P 和 P' 的仿射坐标之差.

按照我们的定义，经过点 P_0 的一维线性流型 S_1 是具有位置向量

$$\overrightarrow{OP} = \mathbf{B} + x_1\mathbf{A}_1$$

的点 P 的轨迹，其中 $\mathbf{B}$ 和 $\mathbf{A}_1$ 是固定的向量 ($\mathbf{A}_1 \neq 0$), 并且 x_1 跑遍所有的实数. 显然， S_1 其实就是经过点 P_0 平行于向量 $\mathbf{A}_1$ 方向的直线 (见第 141 页). 一个二维线性流型或者二维**平面** S_2 是由具有位置向量

$$\overrightarrow{OP} = \mathbf{B} + x_1\mathbf{A}_1 + x_2\mathbf{A}_2$$

的全部点 P 构成的，其中 $\mathbf{B}, \mathbf{A}_1, \mathbf{A}_2$ 是固定的向量 ($\mathbf{A}_1$ 与 $\mathbf{A}_2$ 是无关的), 并且 x_1 和 x_2 跑遍所有的实数. 这 n 维线性空间 S_n 与整个的空间 R^n 是恒等的，因为任何向量 $\mathbf{Y}$ 是相关于 n 个线性无关的向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的 (见第 149 页), 因而任何一点 P 的位置向量都能表示成如下形式:

$$\overrightarrow{OP} = \mathbf{B} + x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n.$$

其次, $n-1$ 维的线性流形能够看作等同于第 149 页定义的超平面. 因为对于 n 维空间中给定的任何 $n-1$ 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{n-1}$, 我们能够找到一个向量 $\mathbf{A}$, 垂直于它们中的每一个 (见第 149 页). 于是，对于

$$\overrightarrow{OP} = \mathbf{B} + x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_{n-1}\mathbf{A}_{n-1}$$

我们有关系

$$\mathbf{A} \cdot \overrightarrow{OP} = \mathbf{B} \cdot \mathbf{A} + x_1\mathbf{A}_1 \cdot \mathbf{A} + x_2\mathbf{A}_2 \cdot \mathbf{A} + \cdots + x_{n-1}\mathbf{A}_{n-1} \cdot \mathbf{A}$$

$= \mathbf{B} \cdot \mathbf{A} =$ 常数,

上式正好是关于点 P 的坐标的一个线性方程.

一般地, 要确定一个向量 $\mathbf{Y}$ 关于基 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的分量 x_i, 即要求形如 (25) 的线性方程组的解. 在一种重要的特殊情形中, 即当基向量作成一个标准正交系时, x_i 能被直接求出来. 对于向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$, 如果它们中每一个的长度为 1 并且任意两个是互相垂直的, 也就是说, 如果

$$\mathbf{A}_i \cdot \mathbf{A}_k = \begin{cases} 1 & 对于\ i = k, \\ 0 & 对于\ i \neq k, \end{cases} \tag{32}$$

我们就称向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 是**标准正交的**(orthonormal). 如果一个向量 $\mathbf{Y}$ 具有以下形式

$$\mathbf{Y} = x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_m\mathbf{A}_m,$$

我们利用垂直关系 (32) 就得到:

$$\mathbf{Y} \cdot \mathbf{A}_i = x_1\mathbf{A}_1 \cdot \mathbf{A}_i + x_2\mathbf{A}_2 \cdot \mathbf{A}_i + \cdots + x_m\mathbf{A}_m \cdot \mathbf{A}_i = x_i \quad (i = 1, 2, \cdots, m). \tag{33}$$

特别, 从 $\mathbf{Y} = \mathbf{0}$ 隐含着 $x_i = 0$ $(i = 1, 2, \cdots, m)$; 所以标准正交向量总是无关的. 公式 (33) 表明, 向量 $\mathbf{Y}$ 关于正交基 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的分量等于向量 $\mathbf{Y}$ 在 $\mathbf{A}_i$ 方向上的分量 $\mathbf{Y} \cdot \mathbf{A}_i$. 由 (22) 式定义的坐标向量 $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n$ 正好作成一个标准正交基, 并且向量 $\mathbf{Y} = (y_1, y_2, \cdots, y_n)$ 关于这个基的分量是

$$\mathbf{Y} \cdot \mathbf{E}_i = y_i.$$

一个标准正交基也可以用如下事实来刻画, 即对于向量的长度和两向量的数量积, 它与原始基 $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n$ 给出相同的公式. 对于任意两个如下形式的向量

$$\begin{aligned} \mathbf{Y} &= x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_m\mathbf{A}_m, \\ \mathbf{Y}' &= x_1'\mathbf{A}_1 + x_2'\mathbf{A}_2 + \cdots + x_m'\mathbf{A}_m, \end{aligned} \tag{34a}$$

我们有

$$
\begin{aligned}
\mathbf{Y}\cdot\mathbf{Y}' =&(x_1\mathbf{A}_1+x_2\mathbf{A}_2+\cdots+x_m\mathbf{A}_m)\\
&\cdot(x_1'\mathbf{A}_1+x_2'\mathbf{A}_2+\cdots+x_m'\mathbf{A}_m)\\
=&x_1\mathbf{A}_1\cdot(x_1'\mathbf{A}_1+x_2'\mathbf{A}_2+\cdots+x_m'\mathbf{A}_m)+\cdots\\
&+x_m\mathbf{A}_m\cdot(x_1'\mathbf{A}_1+x_2'\mathbf{A}_2+\cdots+x_m'\mathbf{A}_m)\\
=&x_1x_1'+x_2x_2'+\cdots+x_mx_m'.^{1)}
\end{aligned}
\tag{34b}
$$

特别当 $\mathbf{Y}=\mathbf{Y}'$ 时，我们得到关于向量 $\mathbf{Y}$ 的长度公式

$$
|\mathbf{Y}|=\sqrt{\mathbf{Y}\cdot\mathbf{Y}}=\sqrt{x_1^2+x_2^2+\cdots+x_m^2}. \tag{34c}
$$

如果通过点 P_0 的 m 维线性流形是由 m 个标准正交向量 $\mathbf{A}_1, \mathbf{A}_2,\cdots,\mathbf{A}_m$ 张成的，则对应的仿射坐标系称为空间 S_m 的一个**笛卡儿坐标系**. 坐标向量 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_m$ 是互相垂直且长度为 1. 具有笛卡儿坐标 $(x_1,x_2,\cdots,x_m)$ 和 $(x_1',x_2',\cdots,x_m')$ 的两个点之间的距离 d 由公式

$$
d=\sqrt{(x_1'-x_1)^2+(x_2'-x_2)^2+\cdots+(x_m'-x_m)^2}
$$

给出. 更一般地，任何以距离概念为基础的几何关系 (如角度，面积，体积) 在任何一个笛卡儿坐标系中都有相同的分析表达式.

b. 矩阵

在 n 维空间中，向量 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_m,\mathbf{Y}$ 之间的这一关系

$$
\mathbf{Y}=x_1\mathbf{A}_1+x_2\mathbf{A}_2+\cdots+x_m\mathbf{A}_m \tag{35a}
$$

1) 没有正交关系，我们只能推出 $\mathbf{Y}\cdot\mathbf{Y}'$ 由一个更复杂的公式给出：

$$
\mathbf{Y}\cdot\mathbf{Y}'=\sum_{ik}c_{ik}x_ix_k,\ \text{其中}\ c_{ik}=\mathbf{A}_i\cdot\mathbf{A}_k.
$$

能够写成一个线性方程组 (见第 148 页 (25) 式)

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= y_1, \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m &= y_2, \\
\cdots\cdots & \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= y_n.
\end{aligned} \tag{35b}
$$

它联系着向量 $\mathbf{Y}$ 在原始坐标系的分量 $y_1, y_2, \cdots, y_n$ 与向量 $\mathbf{Y}$ 关于基向量 $\mathbf{A}_i = (a_{1i}, a_{2i}, \cdots, a_{ni})$, $i = 1, 2, \cdots, m$ 的分量 $x_1, x_2, \cdots, x_m$. 数量 x_i 和 y_i 之间的线性关系 (35b) 完全被这 $n \times m$ 个系数 a_{ji} 所描述. 将这些系数排成一个矩形阵列

$$
\mathbf{a} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}, \tag{36}
$$

叫作**矩阵**(我们将总是用黑体小写字母表示矩阵).

(36) 中的矩阵 $\mathbf{a}$ 有 mn 个 “元素”

$$
a_{ji} \quad (j = 1, 2, \cdots, n;\ i = 1, 2, \cdots, m).
$$

这些元素被排成 m“列 (column)”

$$
\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \cdots, \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}
$$

或者 n“行 (row)”

$$
\begin{aligned}
&(a_{11} \quad a_{12} \quad \cdots \quad a_{1m}), \\
&(a_{21} \quad a_{22} \quad \cdots \quad a_{2m}), \\
&\cdots\cdots
\end{aligned}
$$

$$(a_{n1} \quad a_{n2} \quad \cdots \quad a_{nm}).$$

两个矩阵被认为是相等的，仅当它们的行数相同，列数相同，并且对应元素相等.

矩阵 $\mathbf{a}$ 的列可分别地视为等同于向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的分量的有序组. 我们将常常把由向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的分量构成各列的矩阵 $\mathbf{a}$ 写成这样

$$\mathbf{a} = (\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m). \tag{37}$$

方程组 (35b) 把 n 个数量 $y_1, y_2, \cdots, y_n$ 表示成 m 个数量 $x_1, x_2, \cdots, x_m$ 的线性函数，现在我们能够缩写成简单的符号方程

$$\mathbf{aX} = \mathbf{Y}, \tag{38}$$

其中 $\mathbf{X}$ 表示向量 $(x_1, x_2, \cdots, x_n)$, $\mathbf{Y}$ 表示向量 $(y_1, y_2, \cdots, y_n)$. 如果矩阵 $\mathbf{a}$ 的列向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 是无关的，我们能把 (38) 解释为刻划**一个基的变换**或者是向量坐标系的变换.

这个方程把向量关于子空间 S_m 的基 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的分量 $x_1, x_2, \cdots, x_m$ 联系到该向量关于整个空间 S_n 的基 $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n$ 的分量 $y_1, y_2, \cdots, y_n$. 这可以叫作 (38) 的 "静止" 的解释，其中几何对象 —— 向量 —— 保持固定，而仅仅是参考系转换了.

有另外一个 "活动" 的解释，其中向量看成是变动的，而坐标系看成是不动的. 于是方程 (38) 描写一个**映射**，将 m 维空间中的向量 $(x_1, x_2, \cdots, x_m)$ 映射到 n 维空间中的向量 $(y_1, y_2, \cdots, y_n)$. 方程 (38)，或详细地写出等价方程组 (35b)，所给出的映射叫作 "**线性的**" 或 "**仿射的**". [1)]

1) 在向量的一个仿射映射中，像向量 $\mathbf{Y}$ 的分量 y_i 是原向量 $\mathbf{X}$ 的分量 x_i 的**齐次**线性函数，如公式 (35b) 所示. 如果我们将 $\mathbf{X}, \mathbf{Y}$ 等同于点的**位置向量**，则 (35b) 定义了一个映射，将 m 维空间 R^m 的点 $(x_1, x_2, \cdots, x_m)$ 映为 n 维空间中的点 $(y_1, y_2, \cdots, y_n)$.

例如方程组

$$y_1 = \frac{2}{3}x_1 - \frac{1}{3}x_2, \quad y_2 = -\frac{1}{3}x_1 + \frac{2}{3}x_2$$

$$y_3 = -\frac{1}{3}x_1 - \frac{1}{3}x_2 \tag{38a}$$

对应于矩阵

$$\mathbf{a} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}$$

能够被解释为将平面向量 $\mathbf{X} = (x_1, x_2)$ 映到三维空间中的向量 $\mathbf{Y} = (y_1, y_2, y_3)$ 的一个映射. 此处, 像向量都满足关系

$$y_1 + y_2 + y_3 = 0, \tag{38b}$$

因而它们是垂直于向量 $\mathbf{N} = (1,1,1)$ 的. 把向量 $\mathbf{X}, \mathbf{Y}$ 等同于点的位置向量, 我们就由 (38a) 有一个将 (x_1, x_2) 平面映射到 (y_1, y_2, y_3) 空间中具有方程 (38b) 的平面 π. 在几何上, 点 (y_1, y_2, y_3) 是由点 $(x_1, x_2, 0)$ 垂直地投影到平面 π 而得出的 [1]. 换个说法, 方程 (38a)

用这种方法得到的点的映射是特殊的仿射映射, 它将 R^m 的原点映到 R^n 的原点. 一般的点的仿射映射由**非齐次**线性方程组

$$y_j = \sum_{i=1}^{m} a_{ji}x_i + b_j \quad (j = 1, 2, \cdots, n) \tag{*}$$

给出 (它能从一个把原点映到原点的特殊映射, 再作一个分量为 b_j 的平移得到). 应用映射 (*) 于两个点 $P' = (x_1', x_2', \cdots, x_m')$, $P'' = (x_1'', x_2'', \cdots, x_m'')$, 分别有像点 $Q' = (y_1', y_2', \cdots, y_n')$, $Q'' = (y_1'', y_2'', \cdots, y_n'')$, 我们看到相应的映射将向量 $\overrightarrow{P'P''} = (x_1'' - x_1', x_2'' - x_2', \cdots, x_m'' - x_m')$ 映到向量

$$\overrightarrow{Q'Q''} = (y_1'' - y_1', y_2'' - y_2', \cdots, y_n'' - y_n'),$$

它是由**齐次**方程组 (35b) 给出的.

---

1) 连接 $(x_1, x_2, 0)$ 和 (y_1, y_2, y_3) 的直线平行于平面 π 的法线 N.

能够静止地解释为平面 π 的一个参量表达式，其中 x_1, x_2 充当参量的角色.

不同的矩阵引起不同的线性映射，因为坐标向量

$$\mathbf{E}_1 = (1, 0, \cdots, 0), \quad \mathbf{E}_2 = (0, 1, 0, \cdots, 0), \cdots$$

由 (35b) 映到

$$\mathbf{A}_1 = (a_{11}, a_{21}, \cdots, a_{n1}), \quad \mathbf{A}_2 = (a_{12}, a_{22}, \cdots, a_{n2}), \cdots$$

由此，矩阵 $\mathbf{a}$ 的列向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 正好是坐标向量 $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n$ 的像. 因此矩阵 $\mathbf{a}$ 被映射唯一地确定了.

特别重要的是，把 n 维空间映射到它 **自身之中** 的线性映射 $\mathbf{Y} = \mathbf{aX}$; 这种映射把一个向量 $\mathbf{X} = (x_1, x_2, \cdots, x_n)$ 映到一个向量 $\mathbf{Y} = (y_1, y_2, \cdots, y_n)$, $\mathbf{Y}$ 和 $\mathbf{X}$ 具有相同个数的分量. 这种映射所对应的矩阵行数和列数相同，所以叫作 **方阵**[1]. 用分量写出来，n 行 n 列的方阵 $\mathbf{a}$ 所对应的映射 $\mathbf{Y} = \mathbf{aX}$ 就具有第 151 页 (27) 式的形式. 关于 n 个未知数 n 个线性方程的方程组的解的基础定理 (第 151 页) 现在就能以另一种方式叙述如下:

对于一个方阵 $\mathbf{a}$, **有两个互相排斥的可能情形**:

(1) 对于每一个向量 $\mathbf{X} \neq 0$ 有 $\mathbf{aX} \neq 0$,

(2) 对于某些向量 $\mathbf{X} \neq 0$ 有 $\mathbf{aX} = 0$.

在情形(1), **对于每一个向量** $\mathbf{Y}$ **存在唯一的一个向量** $\mathbf{X}$ 使得 $\mathbf{Y} = \mathbf{aX}$; **在情形**(2), **存在向量** $\mathbf{Y}$, **对于它没有向量** $\mathbf{X}$ **使得** $\mathbf{Y} = \mathbf{aX}$ **成立**.[2]

在情形 (2), 我们称矩阵 $\mathbf{a}$ 为 **奇异的**; 在情形 (1), 称 $\mathbf{a}$ 为 **非奇异的**. 因为方程 $\mathbf{aX} = 0$ 存在着非平凡解 $\mathbf{X}$ 等价于矩阵 $\mathbf{a}$ 的列向

---

1) 一般的行数和列数是任意数的矩阵叫 **矩形** 矩阵.

2) 在情形 (1), 方程 $\mathbf{Y} = \mathbf{aX}$ 表示 n 维空间到它自身上的一个 1–1 映射. 在情形 (2), 映射既不是 1–1 映射，也不是空间自身上的.

量是相关的，我们看到，一个方阵 a 是奇异的，当且仅当它的列向量是相关的．

c．矩阵的运算

习惯上，像 (36) 中那样，用带着两个下标的字母，例如 a_{ji}，来表示矩阵 a 的元素．下标表示元素在矩阵中的位置或地点，第一个下标给出行数，第二个下标给出列数．对于一个具有元素 a_{ji} 的 n 行 m 列的矩阵，下标 j 跑遍 $1, 2, \cdots, n$ 而下标 i 跑遍 $1, 2, \cdots, m$．矩阵 (36) 常常简记作如下形式

$$\mathbf{a} = (a_{ji}),$$

它仅仅显示了矩阵 a 的元素，而没有表示出行数和列数，它们需由前后文关系才能看出[1]．例如在

$$\mathbf{a} = (a_{ji}) = \begin{pmatrix} 1! & 2! & 3! & \cdots & m! \\ 2! & 3! & 4! & \cdots & (m+1)! \\ 3! & 4! & 5! & \cdots & (m+2)! \\ \vdots & \vdots & \vdots & & \vdots \\ n! & (n+1)! & (n+2)! & \cdots & (m+n-1)! \end{pmatrix}$$

中，我们有 $a_{ji} = (i+j-1)!$．

矩阵的加法以及数量对矩阵的乘法是按照对于向量的同一方式来定义的．如果 $\mathbf{a} = (a_{ji})$ 与 $\mathbf{b} = (b_{ji})$ 是"大小"相同的矩阵，就是说，具有相同的行数和列数，我们就把 $\mathbf{a} + \mathbf{b}$ 定义成由对应元素相加而得到的矩阵

$$\mathbf{a} + \mathbf{b} = (a_{ji} + b_{ji}).$$

类似地，对于一个数量 λ 我们把 $\lambda\mathbf{a}$ 定义成用因子 λ 去乘 a 的每个元素所得到的矩阵

$$\lambda\mathbf{a} = (\lambda a_{ji}).$$

---

1) a_{ji} 中的字母 a 表示独立变数 j 和 i 的一个实值函数．这个函数的定义域由 (j, i) 平面上坐标为 $1 \leq j \leq n$ 和 $1 \leq i \leq m$ 的整数的点组成．通常我们将两个独立变量 x, y 的函数写作 $f(x, y)$，于是这里假如用 $a(j, i)$ 代替习惯上的 a_{ji} 就会更为一致．

对于由矩阵决定的向量 X 的映射，可以立即验明有如下的规则：

$$(\mathbf{a}+\mathbf{b})\mathbf{X}=\mathbf{a}\mathbf{X}+\mathbf{b}\mathbf{X},\quad (\lambda\mathbf{a})\mathbf{X}=\lambda(\mathbf{a}\mathbf{X}). \tag{39}$$

更有意义的事实是，“大小”合适的矩阵可以彼此相乘. 两个矩阵 a 与 b 的乘积的一个自然的定义，是由考虑相应的映射 (第一卷第 56 页) 的符号乘积或组合而得到的. 如果 $\mathbf{a}=(a_{ji})$ 是一个具有 m 列和 n 行的矩阵，且 $\mathbf{X}=(x_1,\cdots,x_m)$ 是一个具有 m 个分量的向量，则 a 所确定的映射 $\mathbf{Y}=\mathbf{a}\mathbf{X}$ 将向量 X 映到具有 n 个分量的向量 $\mathbf{Y}=(y_1,\cdots,y_n)$, 其中

$$y_j=\sum_{i=1}^{m}a_{ji}x_i\quad (j=1,2,\cdots,n).$$

如果现在又有矩阵 $\mathbf{b}=(b_{kj})$, 是一个具有 n 列 p 行的矩阵，则映射 $\mathbf{Z}=\mathbf{b}\mathbf{Y}$ 将 Y 映到具有 p 个分量的向量 $\mathbf{Z}=(z_1,\cdots,z_p)$, 其中

$$z_k=\sum_{j=1}^{n}b_{kj}y_j=\sum_{j=1}^{n}\sum_{i=1}^{m}b_{kj}a_{ji}x_i=\sum_{i=1}^{m}c_{ki}x_i,$$

这里

$$c_{ki}=\sum_{j=1}^{n}b_{kj}a_{ji}\quad (k=1,2,\cdots,p,i=1,2,\cdots,m). \tag{40}$$

于是， $\mathbf{Z}=\mathbf{c}\mathbf{X}$, 其中 $\mathbf{c}=\mathbf{b}\mathbf{a}=(c_{ij})$ 是 p 行 m 列的矩阵，它的元素 c_{ki} 由 (40) 式给出. 因此，我们定义矩阵 b 和 a 的乘积 $\mathbf{c}=\mathbf{b}\mathbf{a}$ 为一个矩阵，它的元素 c_{ki} 由 (40) 式给出.

我们注意到，仅仅当矩阵 b 的列数和矩阵 a 的行数相同时，乘积 ba 才有定义. 这相当于以下的一个明显的事实，即两个映射的符号乘积，仅仅在第一个因子的定义域包含了第二个因子的取值范围时，才能够形成. 因此，很可能乘积 ba 有定义，而把因子次序对调后的乘积 ab 却没有定义. 而且，即使两个乘积 ba 与 ab 都有定义时，乘法交换律 $\mathbf{a}\mathbf{b}=\mathbf{b}\mathbf{a}$ 对于矩阵一般地并不成立. 例如

$$\mathbf{a}=\begin{pmatrix}0&1\\-1&0\end{pmatrix},\quad \mathbf{b}=\begin{pmatrix}1&0\\0&-1\end{pmatrix},$$

我们有

$$\mathbf{ab}=\begin{pmatrix}0 & -1\\ -1 & 0\end{pmatrix},\quad \mathbf{ba}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

然而，人们从 (40) 式很容易验证矩阵的乘法遵守结合律和分配律：

$$\mathbf{a}(\mathbf{bc})=(\mathbf{ab})\mathbf{c}, \tag{41a}$$

$$\mathbf{a}(\mathbf{b}+\mathbf{c})=\mathbf{ab}+\mathbf{ac},\quad (\mathbf{a}+\mathbf{b})\mathbf{c}=\mathbf{ac}+\mathbf{bc} \tag{41b}$$

(对于“大小”适当的矩阵而言). 我们可以说，对于矩阵，所有的代数运算是被允许的，只要涉及的乘积有定义并且我们没有交换因子的次序.

由矩阵 $\mathbf{a}$ 所决定的向量的映射，我们已经记作 $\mathbf{Y}=\mathbf{aX}$，能够看作矩阵乘法的特例，只要将向量 $\mathbf{X},\mathbf{Y}$ 写作“列向量”，就是说将它们写作只有一列而分别有 m 行和 n 行的矩阵

$$\mathbf{X}=\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_m\end{pmatrix},\quad \mathbf{Y}=\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}.$$

d. 方阵. 逆阵. 正交阵

在应用上特别重要的是具有行数和列数相等的矩阵，称为**方阵**(行数和列数不相等的矩阵叫**矩形**矩阵). 一个方阵的行数或列数叫作它的**阶**. 任何两个有相同阶数 n 的方阵能够相加或相乘. 特别，对于一个方阵，我们可以作成它的幂

$$\mathbf{a}^2=\mathbf{aa},\quad \mathbf{a}^3=\mathbf{aaa},\cdots.$$

特别，n 阶零方阵 $\mathbf{O}$ 是所有元素都是 0 的方阵，或者说它所有的列向量都是零向量

$$\mathbf{O}=(0,0,\cdots,0). \tag{42a}$$

它具有以下明显的性质

$$\mathbf{a}+\mathbf{0}=\mathbf{0}+\mathbf{a}=\mathbf{a}, \quad \mathbf{a0}=\mathbf{0a}=\mathbf{0} \tag{42b}$$

对于所有的 n 阶方阵 $\mathbf{a}$,

$$\mathbf{0X}=\mathbf{0} \tag{42c}$$

对于所有的具有 n 个分量的向量 $\mathbf{X}$.

其次，单位方阵是用 $\mathbf{e}$ 表示的，是对应着向量 $\mathbf{X}$ 的恒等映射的方阵

$$\mathbf{eX}=\mathbf{X}, \tag{43a}$$

对于所有的向量 $\mathbf{X}$ 成立. 因为对于每一个坐标向量 $\mathbf{E}_k$ 都有

$$\mathbf{eE}_k=\mathbf{E}_k,$$

我们发现单位方阵的列就是坐标向量:

$$\mathbf{e}=(\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_n)=\begin{pmatrix}1 & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}. \tag{43b}$$

容易验证 $\mathbf{e}$ 在方阵的乘法中扮演着“单位”的角色:

$$\mathbf{ae}=\mathbf{ea}=\mathbf{a},$$

对于所有的 n 阶方阵 $\mathbf{a}$ 成立.

如果有 n 阶方阵 $\mathbf{b}$ 和 n 阶方阵 $\mathbf{a}$ 有下式

$$\mathbf{ab}=\mathbf{e}$$

成立，就称 $\mathbf{b}$ 逆于 $\mathbf{a}$. 如果 $\mathbf{b}$ 是逆于 $\mathbf{a}$ 的，则 $\mathbf{a}$ 对应着由 $\mathbf{b}$ 提供的向量映射的逆，就是说，如果 $\mathbf{b}$ 把一个向量 $\mathbf{Y}$ 映射到向量 $\mathbf{X}$ 上 (即如果 $\mathbf{X}=\mathbf{bY}$), 则 $\mathbf{a}$ 把向量 $\mathbf{X}$ 映回到 $\mathbf{Y}$, 因为

$$\mathbf{aX}=\mathbf{abY}=\mathbf{eY}=\mathbf{Y}.$$

更具体地说，如果我们知道方阵 $\mathbf{a}$ 的一个逆方阵 $\mathbf{b}$，则对于任何给定的 $(y_1, y_2, \cdots, y_n) = \mathbf{Y}$，我们能写出线性方程组

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ &\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= y_n \end{aligned} \tag{44}$$

的一组解 $\mathbf{X} = (x_1, x_2, \cdots, x_n)$. 因为有 $\mathbf{abY} = \mathbf{eY} = \mathbf{Y}$. 我们确实得到了原方程的一组解，由 $\mathbf{X} = \mathbf{bY}$ 给出，即由

$$\begin{aligned} x_1 &= b_{11}y_1 + b_{12}y_2 + \cdots + b_{1n}y_n \\ x_2 &= b_{21}y_1 + b_{22}y_2 + \cdots + b_{2n}y_n \\ &\cdots\cdots \\ x_n &= b_{n1}y_1 + b_{n2}y_2 + \cdots + b_{nn}y_n \end{aligned}$$

给出.

每一个异于零的实数 a 都有一个倒数 b 满足 $ab = 1$. 然而却存在异于零方阵的方阵没有逆方阵. 如果 $\mathbf{a}$ 有逆为 $\mathbf{b}$，则方程 $\mathbf{aX} = \mathbf{Y}$ 对于每一个向量 $\mathbf{Y}$ 都有解 $\mathbf{X} = \mathbf{bY}$，因为

$$\mathbf{abY} = \mathbf{eY} = \mathbf{Y}.$$

因此 (见第 165 页) 方阵 $\mathbf{a}$ 必定是非奇异的；就是说 $\mathbf{a}$ 的列向量是无关的. **奇异方阵没有逆方阵**. 对于 $\mathbf{a}$ 的逆方阵 $\mathbf{b}$，条件

$$\mathbf{ab} = \mathbf{e}$$

能写成如下形式

$$\sum_{r=1}^{n} a_{jr} b_{rk} = \mathbf{e}_{jk}, \tag{45}$$

其中 a_{jr}, b_{rk}, e_{jk} 分别表示方阵 $\mathbf{a}, \mathbf{b}, \mathbf{e}$ 的一般元素. 对于固定的 k，我们在 (45) 中有一组关于向量 $\mathbf{B}_k = (b_{1k}, b_{2k}, \cdots, b_{nk})$ 的 n 个线

性方程的方程组，$\mathbf{B}_k$ 为方阵 $\mathbf{b}$ 的第 k 列．如果方阵 $\mathbf{a}$ 是非奇异的，则对于每一个 k，方程组 (45) 都有唯一的解 $\mathbf{B}_k$．因此，非奇异方阵有且仅有一个逆方阵 $\mathbf{b}$．

设 $\mathbf{a}$ 是任一非奇异的方阵，$\mathbf{b}$ 是它的逆，即 $\mathbf{ab}=\mathbf{e}$．考虑任意一个向量 $\mathbf{X}$，且令 $\mathbf{Y}=\mathbf{aX}$．因为 $\mathbf{Z}=\mathbf{X}$ 和 $\mathbf{Z}=\mathbf{bY}$ 都是方程组 $\mathbf{Y}=\mathbf{aZ}$ 的解，又因这解是唯一的，我们必定有

$$\mathbf{bY}=\mathbf{X}, \tag{46}$$

对于每一个向量 $\mathbf{X}$ 成立．因此 (见第 166 页) $\mathbf{a}$ 也是 $\mathbf{b}$ 的逆：

$$\mathbf{ba}=\mathbf{e}.$$

非奇异方阵 $\mathbf{a}$ 的逆方阵常用 $\mathbf{a}^{-1}$ 表示．我们有

$$\mathbf{aa}^{-1}=\mathbf{a}^{-1}\mathbf{a}=\mathbf{e},$$

其中 $\mathbf{e}$ 为单位方阵．对于 b_{rk} 解方程组 (45)，可以推算出逆来．因为单位方阵 $\mathbf{e}$ 的元素 e_{jk}，当 $j\neq k$ 时取值 0；当 $j=k$ 时取值 1，方程组 (45) 表明方阵 $\mathbf{a}$ 的第 j 行与方阵 $\mathbf{a}^{-1}$ 的第 k 列的数量积当 $j\neq k$ 时为 0；当 $j=k$ 时为 1．更进一步，由于

$$\mathbf{a}^{-1}\mathbf{a}=\mathbf{e},$$

我们看到方阵 $\mathbf{a}^{-1}$ 的第 j 行与方阵 $\mathbf{a}$ 的第 k 列的数量积也是当 $j\neq k$ 时为 0；当 $j=k$ 时为 1．

用逆方阵作乘法，使得我们能够用一个非奇异的方阵来除一个方阵之间的等式．例如，方阵等式

$$\mathbf{ab}=\mathbf{c},$$

其中 $\mathbf{a}$ 为一个非奇异的方阵，用 $\mathbf{a}^{-1}$ 从左边乘这等式便能解出 $\mathbf{b}$：

$$\mathbf{a}^{-1}\mathbf{c}=\mathbf{a}^{-1}(\mathbf{ab})=(\mathbf{a}^{-1}\mathbf{a})\mathbf{b}=\mathbf{eb}=\mathbf{b}.$$

类似地，等式

$$\mathbf{ba} = \mathbf{c}$$

导致

$$\mathbf{ca}^{-1} = \mathbf{b}.$$

从欧几里得几何的观点看，最重要的方阵是**正交**(orthogonal)方阵，它对应着从一个笛卡儿坐标系到另一个笛卡儿坐标系的一种过渡，或是保持长度的一种线性变换．我们称 $\mathbf{a}$ 为一正交方阵，如果一个方阵 $\mathbf{a}$ 的列向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 组成一个正交系

$$\mathbf{A}_i \cdot \mathbf{A}_k = \begin{cases} 0, & \text{对于 } i \neq k, \\ 1, & \text{对于 } i = k, \end{cases} \tag{47}$$

(参看第 157 页)．因为组成正交系的向量是无关的，由此可知**正交方阵永远是非奇异的**．对相应于方阵 $\mathbf{a}$ 的向量关系式 $\mathbf{aX} = \mathbf{Y}$ 作“静止”的解释，它便描述了向量关于坐标向量 $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n$ 的分量 $y_1, y_2, \cdots, y_n$ 如何联系到向量关于基向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的分量．对于一个正交方阵 $\mathbf{a}$，一组基 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是由 n 个长度为 1 的正交向量组成的，它们组成一个笛卡儿坐标系，其中距离由常用的表达式 (见第 158 页) 给出．若作“活动”的解释，$\mathbf{Y} = \mathbf{aX}$ 便表示一个线性映射，将坐标向量 $\mathbf{E}_i$ 映成向量 $\mathbf{A}_i$．这个映射将向量

$$\mathbf{X} = (x_1, \cdots, x_n) = x_1\mathbf{E}_1 + x_2\mathbf{E}_2 + \cdots + x_n\mathbf{E}_n$$

映成

$$\begin{aligned} \mathbf{Y} = \mathbf{aX} &= \mathbf{a}(x_1\mathbf{E}_1 + x_2\mathbf{E}_2 + \cdots + x_n\mathbf{E}_n) \\ &= x_1\mathbf{aE}_1 + x_2\mathbf{aE}_2 + \cdots + x_n\mathbf{aE}_n \\ &= x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n. \end{aligned}$$

这种映射保持任何一个向量的长度，因为由 (47) 式有

$$\begin{aligned} |\mathbf{Y}|^2 = \mathbf{Y} \cdot \mathbf{Y} &= (x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n) \\ &\quad \cdot (x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n) \\ &= x_1^2 + x_2^2 + \cdots + x_n^2 = |\mathbf{X}|^2. \end{aligned}$$

更广泛些说，这种映射保持任何两个向量的数量积不变，因此也保持了方向间的夹角，这是容易证明的. 这种保持长度的映射统称为 **正交变换** 或刚体运动. 在二维的情形，它们容易与第一卷 (第 405 页) 讨论过的坐标轴的变换等同起来. 长度为 1 的二维向量 $\mathbf{A}_1$ 能写成 $\mathbf{A}_1 = (\cos\gamma, \sin\gamma)$, 其中 γ 是某一适当的角. 垂直于向量 $\mathbf{A}_1$ 且长度为 1 的向量 $\mathbf{A}_2$ 只有

$$\mathbf{A}_2 = \left(\cos\left(\gamma + \frac{\pi}{2}\right), \sin\left(\gamma + \frac{\pi}{2}\right)\right) = -(\sin\gamma, \cos\gamma)$$

和

$$\mathbf{A}_2 = \left(\cos\left(\gamma - \frac{\pi}{2}\right), \sin\left(\gamma - \frac{\pi}{2}\right)\right) = (\sin\gamma, -\cos\gamma).$$

因此，一般的二阶正交方阵是以下两种形式之一：

$$\mathbf{a} = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \text{ 或 } \mathbf{a} = \begin{pmatrix} \cos\gamma & \sin\gamma \\ \sin\gamma & -\cos\gamma \end{pmatrix}. \tag{48}$$

正交关系 (47) 使人们能立刻写出一个正交方阵 $\mathbf{a}$ 的逆方阵 $\mathbf{a}^{-1}$. 我们取方阵 $\mathbf{a}$ 的列向量 $\mathbf{A}_k$ 作为方阵 $\mathbf{a}^{-1}$ 的行向量；这样 $\mathbf{a}^{-1}$ 的第 k 行和 $\mathbf{a}$ 的第 j 列的数量积当 $j \neq k$ 时为 0, 当 $j = k$ 时为 1, 满足了 $\mathbf{a}^{-1}\mathbf{a} = \mathbf{e}$ 的要求. 一般地，对于任意一个矩阵 $\mathbf{a} = (a_{jk})$, 交换它的行和列，得到矩阵 $\mathbf{a}$ 的转置矩阵 $\mathbf{a}^T = (b_{jk})$. 更明确些有 $b_{jk} = a_{kj}$ [1]. 对于一个正交方阵，我们简单地有

$$\mathbf{a}^{-1} = \mathbf{a}^T. \tag{49}$$

例如

$$\begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}^{-1} = \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix}.$$

由于 (46), 我们能将关系式 (49) 写成

$$\mathbf{a}^T\mathbf{a} = \mathbf{e}, \quad \mathbf{a}\mathbf{a}^T = \mathbf{e}.$$

---

1) 将 $\mathbf{a}$ 写成一个矩形阵列，定义 $\mathbf{a}$ 的"主对角线"是从左上角向右下角斜率为 -1 的直线，它包含元素 $a_{11}, a_{22}, a_{33}, \cdots$. 将矩阵 $\mathbf{a}$ 以主对角线为准来"反射"，便得到矩阵 $\mathbf{a}$ 的转置矩阵.

这第二个关系表明，在一个正交方阵中，第 j 行和第 k 行的数量积当 $j \neq k$ 时为 0，当 $j = k$ 时为 1. 因此**一个正交方阵的行向量也作成一个正交系**.

练 习 2.2

1. 对下列每个情形写出经过点 P 由向量 $\mathbf{A}_k$ 张成的空间.

(a) $P = (-1,2,1)$; $\mathbf{A}_1 = (4,0,3)$.

(b) $P = (2,1,-4)$; $\mathbf{A}_1 = (3,-2,1)$, $\mathbf{A}_2 = (1,0,-1)$.

(c) $P = (2,1,-4,2)$; $\mathbf{A}_1 = (3,-2,1,2)$, $\mathbf{A}_2 = (1,0,-1,2)$.

2. 证明 $\mathbf{E}_1 = (2/3,2/3,-1/3)$、$\mathbf{E}_2 = (1/\sqrt{2},-1/\sqrt{2},0)$, $\mathbf{E}_3 = (\sqrt{2}/6, \sqrt{2}/6, 2\sqrt{2}/3)$ 作成一个正交基，并写出下列给定的向量在这组基下的表达式：

(a) $\mathbf{A}_1 = (\sqrt{2},\sqrt{2},\sqrt{2})$,

(b) $\mathbf{A}_2 = (3,-3,3)$,

(c) $\mathbf{A}_3 = (1,0,0)$.

3. 给定了线性无关的向量 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_m$，作出互相垂直的单位向量 $\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_m$ 具有如下性质，即 $\mathbf{E}_k$ 是 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_k$ 的线性组合 (对于 $k = 1,2,\cdots,m$).

4. 从练习第 3 题的结果，证明线性无关的基本定理.

5. 求直线到点 $P = (x_0,y_0,z_0)$ 的距离，直线由下式给出：

$$x = at + b, \quad y = ct + d, \quad z = et + f.$$

(提示：找 P 点到直线的垂足.)

6. 以下方程组是否有一组非零解？

$$x + 2y + 3z = 0,$$

$$2x + 3y + z = 0,$$

$$3x + y + 2z = 0.$$

7. 写出向量 (a_1, a_2, a_3) 关于基 $\mathbf{A}_1 = (1,2,3)$, $\mathbf{A}_2 = (2,3,1)$, $\mathbf{A}_3 = (3,1,2)$ 的表达式.

8. 确定从关于基 $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ 的笛卡儿坐标系到关于练习第 7 题中的基 $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ 的仿射坐标系的变换的矩阵.

9. 如果方阵 $\mathbf{a}$ 是奇异的，证明存在一个向量 $\mathbf{Y}$, 对于它 $\mathbf{Y} = \mathbf{aX}$ 无解.

10. 对于矩阵

$$\mathbf{a} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -2 & 1 & 0 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{pmatrix}$$

求乘积 $\mathbf{ab}$ 和 $\mathbf{ba}$.

11. 找出 2 阶方阵

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

存在逆方阵的条件. 当逆方阵存在时把它写出来.

12. 证明只有一个单位方阵.

13. 设方阵 $\mathbf{a}, \mathbf{b}$ 都是非奇异的，求出 $\mathbf{ab}$ 的逆方阵.

14. 有时人们把一个 n 阶奇异方阵定义为将 n 维空间映射成一个低维空间的矩阵. 证明这个定义与前面给出的定义是等价的.

15. 说明 (48) 中矩阵的几何意义.

16. 证明方阵 $\mathbf{a}$ 是正交的当且仅当 $\mathbf{a}^T = \mathbf{a}^{-1}$.

17. 证明乘积 $\mathbf{ab}$ 的转置矩阵是转置矩阵交换次序后的乘积 $\mathbf{b}^T\mathbf{a}^T$.

18. 证明正交方阵的乘积是正交的.

19. 证明正交方阵对应的映射保持数量积；就是说，如果 $\mathbf{a}$ 是正交的，则 $(\mathbf{aX}) \cdot (\mathbf{aY}) = \mathbf{X} \cdot \mathbf{Y}$.

20. 证明任何保持长度的线性映射的方阵是正交的.

21. 证明一个仿射变换将质点组的质量中心变到像点组的质量中心.

2.3 行 列 式

a. 二阶与三阶行列式

数学分析包括多维空间中非线性映射的研究. 然而作这样的研究之前必须先有一个关于线性映射 $\mathbf{Y}=\mathbf{aX}$ 的研究, 其中 $\mathbf{X}$ 和 $\mathbf{Y}$ 是向量而 $\mathbf{a}$ 是矩阵. 尤其具有根本重要性的是分析一个线性映射的逆的构造, 这实际上就是分析 n 个线性方程

$$\begin{aligned}
a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n&=y_1,\\
a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n&=y_2,\\
&\cdots\cdots\\
a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n&=y_n
\end{aligned}\tag{50}$$

对于 n 个未知量 $x_1,x_2,\cdots,x_n$ 的解的构造.

解 n 个未知量的 n 个线性方程的过程引出某种有很多项的代数表达式, 叫行列式. 行列式的显式定义和性质在开始时显得有些神秘. 到了我们把行列式的定义建立在 n 维空间中 n 个向量的多重线性交替型这样一个单独的性质上时, 这些神秘性将逐渐消失. 从这个概念性的途径出发, 行列式的所有重要性质都能容易地导出来. 我们将在本书后面几章看到, 在推广到高维的微积分中行列式是极端重要的.

对于开始很少几个 n 值, 写出方程组 (50) 的显式解是有教益的. 对于 $n=1$, 我们有这单个的方程

$$a_{11}x_1=y_1,$$

其解为

$$x_1=y_1/a_{11}.\tag{50a}$$

对于 $n=2$, 我们有方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2=y_1,\\ a_{21}x_1+a_{22}x_2=y_2.\end{cases}$$

用 a_{22} 乘第一个方程，用 a_{12} 乘第二个方程，然后作减法消去 x_2，就得到关于 x_1 的单个方程；类似地，可用 a_{21} 乘第一个方程，用 a_{11} 乘第二个方程，然后作减法消去 x_1. 这样就得到关于 x_1, x_2 的表达式

$$x_1 = \frac{a_{22}y_1 - a_{12}y_2}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{a_{11}y_2 - a_{21}y_1}{a_{11}a_{22} - a_{12}a_{21}}. \tag{50b}$$

对于 $n = 3$, 我们有方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = y_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = y_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = y_3. \end{cases} \tag{50c}$$

我们能够把这个方程组化为关于 x_1, x_2 的两个方程，譬如这样来消去 x_3: 用 a_{13}/a_{23} 乘第二个方程，并从第一个方程减去它，再用 a_{13}/a_{33} 乘第三个方程，也从第一个方程减去它. 这样得到的只含 x_1, x_2 的两个方程可用上面的方法解出来. 通过一些代数运算之后，我们得到

$$\begin{aligned} x_1 =&(a_{22}a_{33}y_1 + a_{12}a_{23}y_2 + a_{13}a_{32}y_2 - a_{13}a_{22}y_3 - a_{23}a_{32}y_1 - a_{12}a_{33}y_2) \\ &/(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} \\ &- a_{12}a_{21}a_{33}), \end{aligned} \tag{50d}$$

对于 x_2 和 x_3 也有类似的公式. 对于 $n = 4$, 计算变得很庞大，显然只有系统性的研究方法才能够产生整齐有序的结果.

我们注意到，在每个情形，解 x_i 都取商的形式，其中分母仅仅是系数 a_{ji} 的函数，就是说是矩阵 $\mathbf{a} = (a_{ji})$ 的函数. 对于 $n = 1$, 这个函数就仅仅是系数 a_{11} 本身. 对于 $n = 2$, 由矩阵

$$\mathbf{a} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

的元素所形成的分母

$$a_{11}a_{22} - a_{12}a_{21}$$

叫作**矩阵 a 的行列式**，记作

$$a_{11}a_{22}-a_{12}a_{21}=\det(\mathbf{a})=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}. \tag{51a}$$

显然 (50b) 的分子也能写成行列式，给出表达式

$$x_1=\frac{\begin{vmatrix} y_1 & a_{12} \\ y_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}};\quad x_2=\frac{\begin{vmatrix} a_{11} & y_1 \\ a_{21} & y_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \tag{51b}$$

当然，这些公式，只有当分母中的行列式不等于零时才有意义.

公式 (50d) 启发我们引进三阶矩阵

$$\mathbf{a}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

的行列式，即表达式

$$a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}$$
$$=\det(\mathbf{a})=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \tag{52a}$$

这样一个三阶行列式的形成规律能够用容易记忆的“对角线规则”来表示 (图 2.5a). 我们在第三列后面重复前两列；做出每一个对角

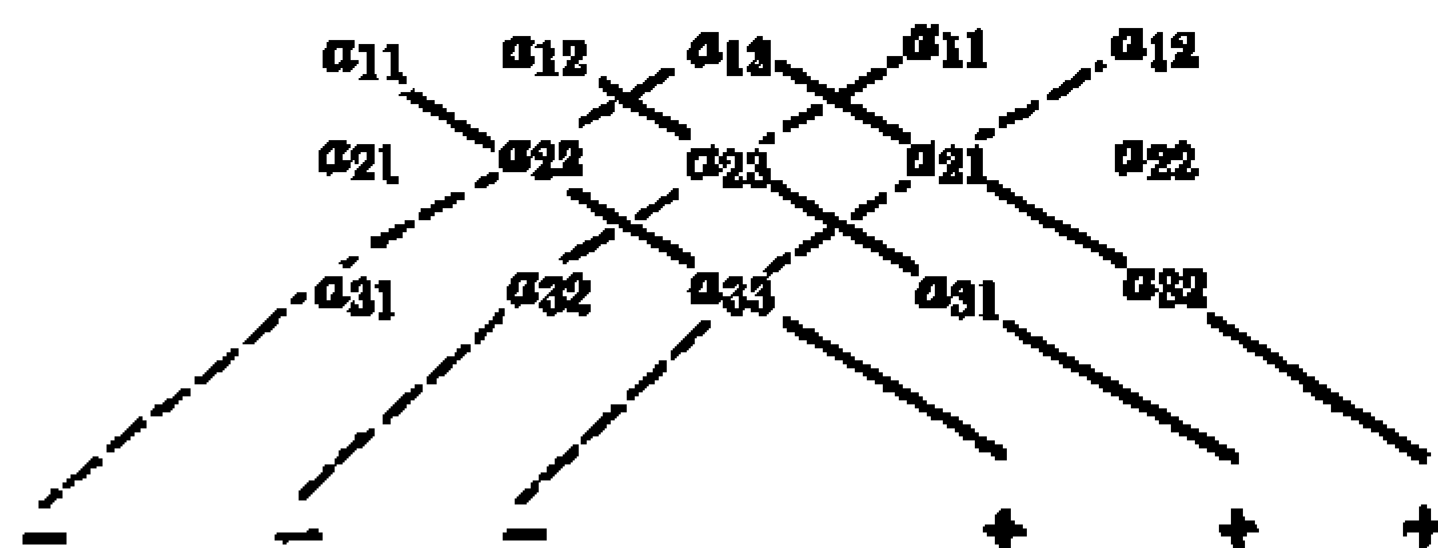


图 2.5a

线上三元组的乘积，对于联系着向右下斜的线的乘积乘以 +1，向左下斜的乘以 −1；然后相加. (这个规则仅仅对于三阶行列式成立！)

利用三阶行列式，我们能把方程组 (50c) 的解写成以下简明的形式：

$$x_1 = \frac{\begin{vmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & y_1 & a_{13} \\ a_{21} & y_2 & a_{23} \\ a_{31} & y_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}},$$

$$x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & y_1 \\ a_{21} & a_{22} & y_2 \\ a_{31} & a_{32} & y_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}.$$

在 (50a) 的基础上，我们类似地定义一阶矩阵

$$\mathbf{a} = (a_{11})$$

的行列式为

$$a_{11} = \det(\mathbf{a}).$$

于是我们看到，对于 $n = 1, 2, 3$ 的每个情形，方程组 (50) 的解 $(x_1, \cdots, x_n)$ 可以描述如下（"克拉默法则"）：**每个未知数** x_i **是两个行列式的商．分母是矩阵** $\mathbf{a} = (a_{jk})$ **的行列式；分子也是一个矩阵的行列式，这矩阵是将矩阵** $\mathbf{a}$ **的第** i **列换成方程组右边的量** $y_1, y_2, \cdots, y_n$ **所得到的矩阵．**

b. 向量的线性型与多线性型

为了定义高阶行列式，以及表述它们的主要性质，必须要使用某些一般的代数概念．

我们可以把 n 个自变量 $a_1, a_2, \cdots, a_n$ 的一个函数 $f(a_1, a_2, \cdots, a_n$ 看作**向量** $\mathbf{A} = (a_1, a_2, \cdots, a_n)$ **的函数**，并记作 $f(\mathbf{A})$．我们称 f 是向量 $\mathbf{A}$ 的一个**线性型**，如果对于任何两个向量 $\mathbf{A}$ 和 $\mathbf{B}$ 都有

$$f(\mathbf{A} + \mathbf{B}) = f(\mathbf{A}) + f(\mathbf{B}), \tag{53a}$$

并且对于任何向量 $\mathbf{A}$ 和数量 λ 都有

$$f(\lambda\mathbf{A}) = \lambda f(\mathbf{A}). \tag{53b}$$

这两个规则 (53a,b) 能够合并为

$$f(\lambda\mathbf{A} + \mu\mathbf{B}) = \lambda f(\mathbf{A}) + \mu f(\mathbf{B}), \tag{54a}$$

对于任何向量 $\mathbf{A}$、$\mathbf{B}$ 和数量 λ, μ 都成立. 写得详细些, 规则 (54a) 就成为

$$\begin{aligned} &f(\lambda a_1 + \mu b_1, \lambda a_2 + \mu b_2, \cdots, \lambda a_n + \mu b_n) \\ =&\lambda f(a_1, a_2, \cdots, a_n) + \mu f(b_1, b_2, \cdots, b_n). \end{aligned} \tag{54b}$$

例如函数

$$f(\mathbf{A}) = 3a_2 - 27a_3$$

是一个线性型, 而

$$f(\mathbf{A}) = |\mathbf{A}| = \sqrt{a_1^2 + \cdots + a_n^2}$$

不是线性型.

关系式 (54a) 直接隐含着线性型的更一般的规则

$$\begin{aligned} &f(\lambda\mathbf{A}_1 + \lambda_2\mathbf{A}_2 + \cdots + \lambda_m\mathbf{A}_m) \\ =&\lambda_1 f(\mathbf{A}_1) + \lambda_2 f(\mathbf{A}_2) + \cdots + \lambda_m f(\mathbf{A}_m) \end{aligned} \tag{54c}$$

对于任何 m 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 和数量 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 都成立. 这个规则对于向量 $\mathbf{A}$ 的最一般的线性型给出一个显式表达式. 利用坐标向量 $\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n$, 我们由 (2b) 便有

$$\mathbf{A} = (a_1, a_2, \cdots, a_n) = a_1\mathbf{E}_1 + a_2\mathbf{E}_2 + \cdots + a_n\mathbf{E}_n,$$

对于任何向量 $\mathbf{A}$ 成立. 因此, 根据 (54c), f 有如下形式

$$f(\mathbf{A}) = a_1 f(\mathbf{E}_1) + a_2 f(\mathbf{E}_2) + \cdots + a_n f(\mathbf{E}_n)$$

$$= c_1a_1 + c_2a_2 + \cdots + c_na_n, \tag{55a}$$

其中 c_i 为常数

$$c_i = f(\mathbf{E}_i). \tag{55b}$$

将系数 c_i 结合成向量 $\mathbf{C} = (c_1, c_2, \cdots, c_n)$, 我们有

$$f(\mathbf{A}) = \mathbf{C} \cdot \mathbf{A}. \tag{55c}$$

向量 $\mathbf{A}$ **的最一般的线性型是向量** $\mathbf{A}$ **与一适当的常向量** $\mathbf{C}$ **的数量积.**

两个向量 $\mathbf{A} = (a_1, a_2, \cdots, a_n)$ 和 $\mathbf{B} = (b_1, b_2, \cdots, b_n)$ 的一个函数 $f(\mathbf{A}, \mathbf{B})$ 称为向量 $\mathbf{A}, \mathbf{B}$ 的一个**双线性型**, 如果当 $\mathbf{B}$ 固定时它是 $\mathbf{A}$ 的线性型, 而当 $\mathbf{A}$ 固定时它是 $\mathbf{B}$ 的线性型. 这就是说, 对于任何向量 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ 和数量 λ, μ 都有

$$f(\lambda\mathbf{A} + \mu\mathbf{B}, \mathbf{C}) = \lambda f(\mathbf{A}, \mathbf{C}) + \mu f(\mathbf{B}, \mathbf{C}), \tag{56a}$$

$$f(\mathbf{A}, \lambda\mathbf{B} + \mu\mathbf{C}) = \lambda f(\mathbf{A}, \mathbf{B}) + \mu f(\mathbf{A}, \mathbf{C}). \tag{56b}$$

双线性型的最简单的例子是数量积

$$f(\mathbf{A}, \mathbf{B}) = \mathbf{A} \cdot \mathbf{B}.$$

在这个例子里, 规则 (56a,b) 正好化成对于数量积的结合律和分配律, 见第 142 页 (15b,c).

从 (56a, b) 我们得到更一般的关系式

$$f(\alpha\mathbf{A} + \beta\mathbf{B}, \gamma\mathbf{C} + \delta\mathbf{D}) = \alpha f(\mathbf{A}, \gamma\mathbf{C} + \delta\mathbf{D}) + \beta f(\mathbf{B}, \gamma\mathbf{C} + \delta\mathbf{D})$$

(56

$$= \alpha\gamma f(\mathbf{A}, \mathbf{C}) + \alpha\delta f(\mathbf{A}, \mathbf{D}) + \beta\gamma f(\mathbf{B}, \mathbf{C}) + \beta\delta f(\mathbf{B}, \mathbf{D}$$

因此我们能对双线性型进行运算, 就像运算通常的表达式的乘积一样. 再利用向量 $\mathbf{A}, \mathbf{B}$ 的分解式

$$\mathbf{A} = (a_1, a_2, \cdots, a_n) = a_1\mathbf{E}_1 + a_2\mathbf{E}_2 + \cdots + a_n\mathbf{E}_n,$$

$$\mathbf{B}=(b_1,b_2,\cdots,b_n)=b_1\mathbf{E}_1+b_2\mathbf{E}_2+\cdots+b_n\mathbf{E}_n,$$

我们就得到公式

$$\begin{aligned}f(\mathbf{A},\mathbf{B})&=f(a_1\mathbf{E}_1+a_2\mathbf{E}_2+\cdots+a_n\mathbf{E}_n,\\&\qquad b_1\mathbf{E}_1+b_2\mathbf{E}_2+\cdots+b_n\mathbf{E}_n)\\&=\sum_{j,k=1}^{n}a_jb_kf(\mathbf{E}_j,\mathbf{E}_k).\end{aligned}$$

因此，向量 $\mathbf{A},\mathbf{B}$ 的最一般的双线性型具有这个形式：

$$f(\mathbf{A},\mathbf{B})=\sum_{j,k=1}^{n}c_{jk}a_jb_k, \tag{57a}$$

带有常数系统

$$c_{jk}=f(\mathbf{E}_j,\mathbf{E}_k). \tag{57b}$$

在 $\mathbf{B}=\mathbf{A}$ 的情形，双线性型 f 转化为**二次型**

$$f(\mathbf{A},\mathbf{A})=\sum_{j,k=1}^{n}c_{jk}a_ja_k. \tag{57c}$$

用类似的方式，人们定义三个向量 $\mathbf{A},\mathbf{B},\mathbf{C}$ 的三线性型 $f(\mathbf{A},\mathbf{B},\mathbf{C})$，作为一个函数，它分别对于每个向量都是线性型．完全与前面一样，最一般的三线性型具有这个形式：

$$f(\mathbf{A},\mathbf{B},\mathbf{C})=\sum_{j,k,r=1}^{n}c_{jkr}a_jb_kc_r, \tag{58a}$$

其中

$$c_{jkr}=f(\mathbf{E}_j,\mathbf{E}_k,\mathbf{E}_r). \tag{58b}$$

更一般地，任意 m 个向量的**多线性型**都能够用一种明显的方式来定义．由于我们不再对不同的向量使用不同的字母，这不过是

引进一种新足标的记号问题．我们用 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 表示向量，并通过

$$\mathbf{A}_1 = (a_{11}, a_{21}, \cdots, a_{n1}), \quad \mathbf{A}_2 = (a_{12}, a_{22}, \cdots, a_{n2}), \cdots,$$
$$\mathbf{A}_m = (a_{1m}, a_{2m}, \cdots, a_{nm})$$

引进它们的分量 a_{jk}. 函数 f 是关于向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的一个多线性型 $f(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m)$，如果它对于每一个向量而言当其他向量固定时都是线性型．我们也可以把 f 看作具有列向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的矩阵

$$\mathbf{a} = (\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m) = (a_{jk})$$

的函数．类似于 (58a)，向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 的最一般的多线性型具有形式

$$f(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m) = \sum_{\substack{j_1, j_2, \cdots, j_m \\ =1,2,\cdots,n}} c_{j_1 j_2 \cdots j_m} a_{j_1 1} a_{j_2 2} \cdots a_{j_m m}, \qquad (59a)$$

其中[1]

$$c_{j_1 j_2 \cdots j_m} = f(\mathbf{E}_{j_1}, \mathbf{E}_{j_2}, \cdots, \mathbf{E}_{j_m}). \qquad (59b)$$

c. 多线性交替型．行列式的定义

公式 (51a) 和 (52a) 所定义的二阶和三阶行列式是特殊的多线性型．第 175 页 (51a) 的二阶行列式是二维向量

$$\mathbf{A}_1 = (a_{11}, a_{21}), \quad \mathbf{A}_2 = (a_{12}, a_{22}) \qquad (60a)$$

的双线性型；(52a) 的三阶行列式是三维向量

$$\mathbf{A}_1 = (a_{11}, a_{21}, a_{31}), \quad \mathbf{A}_2 = (a_{12}, a_{22}, a_{32}), \mathbf{A}_3 = (a_{13}, a_{23}, a_{33}) \qquad (60b)$$

1) 在这个公式中，应用下标的下标是很麻烦的．这里 $j_1, j_2, \cdots, j_m$ 表示从数的集合 $1, 2, \cdots, n$ 选出的任意 m 个数的组合．这样的组合也能当作一个函数 $j(k)$，它的定义域是数的集合 $k = 1, 2, \cdots, m$，它的取值范围是数的集合 $j = 1, 2, \cdots, n$．这些组合或函数中的任何一个给出公式 (59a) 中的一项．

的三线性型. (行列式分别对于每一个向量的线性性质, 可以通过检查下面的事实得到, 即在行列式的展开式的每一个乘积中, 关于给定的第二个下标正好包含一个因子.) 把行列式与其他多线性型区别开来的附加特性是它们的"交替"特性.

几个变元 (它们可以是向量也可以是数量) 的一个函数, 如果当我们交换任何两个变量时它仅仅改变符号, 就说这个函数是交替的. 数量变元的交替函数的例子是

$$\phi(x, y) = y - x, \tag{61a}$$

$$\phi(x, y, z) = (z - y)(z - x)(y - x). \tag{61b}$$

两个 n 维向量 $\mathbf{A}_1, \mathbf{A}_2$ 的一个函数 f 是交替的, 如果对于任意两个向量 $\mathbf{A}_1, \mathbf{A}_2$ 都有

$$f(\mathbf{A}_1, \mathbf{A}_2) = -f(\mathbf{A}_2, \mathbf{A}_1).$$

这特别隐含着, 对于 $\mathbf{A}_1 = \mathbf{A}_2 = \mathbf{A}$ 有

$$f(\mathbf{A}, \mathbf{A}) = 0.$$

设 $n = 2$ 并且 f 是 (60a) 所给定的关于向量 $\mathbf{A}_1, \mathbf{A}_2$ 的一个交替函数, 它也是一个双线性型. 于是

$$f(\mathbf{E}_1, \mathbf{E}_1) = f(\mathbf{E}_2, \mathbf{E}_2) = 0, \quad f(\mathbf{E}_2, \mathbf{E}_1) = -f(\mathbf{E}_1, \mathbf{E}_2).$$

由 (57a, b) 得到:

$$\begin{aligned} f(\mathbf{A}_1, \mathbf{A}_2) &= f(a_{11}\mathbf{E}_1 + a_{21}\mathbf{E}_2, a_{12}\mathbf{E}_1 + a_{22}\mathbf{E}_2) = c(a_{11}a_{22} - a_{12}a_{21}) \\ &= c\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = c \det(\mathbf{A}_1, \mathbf{A}_2), \end{aligned} \tag{62a}$$

其中常数 c 的值为

$$c = f(\mathbf{E}_1, \mathbf{E}_2). \tag{62b}$$

所以，在二维空间中，两个向量 $\mathbf{A}_1, \mathbf{A}_2$ 的每一个双线性交替型同具有列向量 $\mathbf{A}_1, \mathbf{A}_2$ 的方阵的行列式仅仅相差一个常数因子 c.

更一般地，两个 n 维向量 $\mathbf{A}_1, \mathbf{A}_2$ 的任何一个双线性交替型都能写成

$$f(\mathbf{A}_1, \mathbf{A}_2) = \sum_{j,k=1}^{n} c_{jk} a_{j1} a_{k2},$$

其中

$$c_{jk} = -c_{kj}, \quad c_{jj} = 0.$$

将下标仅仅相差一个对换的项组合起来，就能将函数 f 表示成二阶行列式的线性组合:

$$\begin{aligned} f(\mathbf{A}_1, \mathbf{A}_2) &= \sum_{\substack{j,k=1 \\ j<k}}^{n} c_{jk}(a_{j1}a_{k2} - a_{k1}a_{j2}) \\ &= \sum_{\substack{j,k=1 \\ j<k}}^{n} c_{jk} \begin{vmatrix} a_{j1} & a_{k1} \\ a_{j2} & a_{k2} \end{vmatrix}. \end{aligned} \tag{62c}$$

对于三个向量的交替函数 f，我们有关系式

$$\begin{aligned} f(\mathbf{A}, \mathbf{B}, \mathbf{C}) &= -f(\mathbf{B}, \mathbf{A}, \mathbf{C}) = -f(\mathbf{A}, \mathbf{C}, \mathbf{B}) \\ &= -f(\mathbf{C}, \mathbf{B}, \mathbf{A}). \end{aligned} \tag{63a}$$

由此又有

$$f(\mathbf{A}, \mathbf{B}, \mathbf{C}) = f(\mathbf{B}, \mathbf{C}, \mathbf{A}) = f(\mathbf{C}, \mathbf{A}, \mathbf{B}). \tag{63b}$$

特别是，当 f 的变元有两个相等时，它等于 0. 设 $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ 是 (60b) 所给定的三维向量. 由 (58a, b)，向量 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ 的一般三线性交替型是

$$f(\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3) = \sum_{j,k,r=1}^{3} c_{jkr} a_{j1} a_{k2} a_{r3}.$$

这里, 应用 (63a, b) 便有

$$c_{jkr} = f(\mathbf{E}_j, \mathbf{E}_k, \mathbf{E}_r) = \varepsilon_{jkr} f(\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3),$$

并且, 如果 j, k, r 中有两个相等, 便有 $\varepsilon_{jkr} = 0$, 此外便有

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1, \quad \varepsilon_{213} = \varepsilon_{132} = \varepsilon_{321} = -1. \tag{64a}$$

应用 (61b) 中的函数 ϕ 的这个性质, 即当它的任何两个变元交换位置时函数值变号, 我们就得到关于 ε_{jkr} 的简明表达式

$$\varepsilon_{jkr} = \operatorname{sgn} \phi(j, k, r) = \operatorname{sgn}(r-k)(r-j)(k-j). \tag{64b}$$

与第 163 页公式 (52a) 的三阶行列式的算式相比较便得到

$$f(\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3) = c \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \tag{64c}$$

其中 $c = f(\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3)$ 是一个常数. 于是我们得到与二维情形同样的结果: 三维向量 $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ 的最一般的三线性交替型同具有列向量 $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ 的方阵的行列式仅仅相差一个常数因子 c. 于是很显然, 以 $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ 作为列向量的方阵的三阶行列式就是那唯一确定的 "关于" 向量 $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ 的三线性交替型, 它当 $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ 分别为坐标向量 $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ 时取值为 1[1].

现在已很清楚, 我们应该怎样来定义高阶行列式了. 设 $\mathbf{a}$ 是具有列向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的方阵

$$\mathbf{a} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{65a}$$

设 f 是 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的一个多线性交替型. 于是 f 由 (59a) 给出, 其中系数 $c_{j_1 j_2 \cdots j_n}$ 有如下形式

$$c_{j_1 j_2 \cdots j_n} = f(\mathbf{E}_{j_1}, \mathbf{E}_{j_2}, \cdots, \mathbf{E}_{j_n}), \tag{65b}$$

1) 最后一个条件表示单位方阵的行列式的值为 1.

并且当交换 $j_1, j_2, \cdots, j_n$ 中任意两个数时，它改变符号．现在用 $\phi(x_1, x_2, \cdots, x_n)$ 表示乘积

$$\begin{aligned}\phi(x_1, x_2, \cdots, x_n) =& (x_n - x_{n-1})(x_n - x_{n-2}) \cdots (x_n - x_2)(x_n - x_1) \\ & \cdot (x_{n-1} - x_{n-2}) \cdots (x_{n-1} - x_2)(x_{n-1} - x_1) \\ & \cdots\cdots \\ & \cdot (x_3 - x_2)(x_3 - x_1)(x_2 - x_1) \\ =& \prod_{\substack{j,k=1,2,\cdots,n \\ j<k}} (x_k - x_j). \end{aligned} \tag{65c}$$

容易看出 ϕ 是数量 $x_1, x_2, \cdots, x_n$ 的一个交替函数，仅仅当这些数中两个相等时它才为零．于是

$$\varepsilon_{j_1 j_2 \cdots j_n} = \operatorname{sgn} \phi(j_1, j_2, \cdots, j_n) \tag{65d}$$

是 $j_1, j_2, \cdots, j_n$ 的一个交替函数，它仅仅取值 $+1, 0, -1$. 对于取值在 $1, 2, \cdots, n$ 的 $j_1, j_2, \cdots, j_n$，除非数字 $j_1, j_2, \cdots, j_n$ 两两不同，就是说，除非它们构成数字 $1, 2, \cdots, n$ 的一个**排列**，我们就有 $\varepsilon_{j_1 j_2 \cdots j_n} = 0$. 如果 $\varepsilon_{j_1 j_2 \cdots j_n} = +1$，人们就称 $j_1, j_2, \cdots, j_n$ 是 $1, 2, \cdots, n$ 的一个**偶排列**；如果 $\varepsilon_{j_1 j_2 \cdots j_n} = -1$ 就称作是一个**奇排列**．一个偶排列能够通过偶数次两个元素的对换重排成 $1, 2, \cdots, n$ 的顺序；一个奇排列则是通过奇数次的这种对换．

由 (65b)，显然有

$$c_{j_1 j_2 \cdots j_n} = \varepsilon_{j_1 j_2 \cdots j_n} f(\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n). \tag{65e}$$

我们定义 (65a) 中方阵 $\mathbf{a}$ 的行列式为

$$\det(\mathbf{a}) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \sum_{j_1 \cdots j_n = 1}^{n} \varepsilon_{j_1 j_2 \cdots j_n} a_{j_1 1} a_{j_2 2} \cdots a_{j_n n}. \tag{66a}$$

于是有结果: n 个 n 维向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的最一般的多线性交替型 f 同具有列向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的方阵的行列式仅仅相差一个常数因子 $c = f(\mathbf{E}_1, \mathbf{E}_2, \cdots, \mathbf{E}_n)$.

d. 行列式的主要性质

公式 (66a) 给出了 n 阶行列式表成它的 n^2 个元素 a_{jk} 的显式展开式. 只计算 $\varepsilon_{j_1 j_2 \cdots j_n}$ 不为零的项, 行列式便是由 $n!$ 个项构成的关于 a_{jk} 的一个 n 次型. 每一项 (先不管系数 $\varepsilon_{j_1 j_2 \cdots j_n} = \pm 1$) 是 n 个元素的乘积, 它们分别属于不同的行和列. 原则上, 对于元素的任何给定值, 都能够由展开式计算行列式的值; 但在实际上, 为了数值计算, 公式要用到的项太多了 (五阶行列式有 120 项, 10 阶行列式有 3,628,800 项), 因而人们设计了各种计算行列式的更有效的方法.

行列式的基本性质已经包含在我们的定义中了, 这就是把它作为 n 维空间中 n 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 的多线性交替型. 如果 $\mathbf{a}$ 表示以这些向量为列向量的方阵, 我们就把这个行列式记为

$$\det(\mathbf{a}) = \det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n).$$

于是, 立即推知, 如果交换方阵 $\mathbf{a}$ 的任意两列, $\mathbf{a}$ 的行列式就改变符号; 特别是, 当方阵 $\mathbf{a}$ 有两列全同时, 它的行列式等于零. 利用行列式分别对于每一个列向量的线性性质, 我们发现, 用一个因子 λ 乘方阵 $\mathbf{a}$ 的一个列, 效果相当于用 λ 乘方阵 $\mathbf{a}$ 的行列式[1]. 例如

$$\det(\lambda \mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n) = \lambda \det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n). \tag{67a}$$

1) 用因子 λ 乘 n 阶方阵 $\mathbf{A}$ 的所有的元素, 等价于用 λ 乘它的 n 个列中的每一列, 因此产生的结果是用 λ^n 乘 $\mathbf{a}$ 的行列式. 即 $\det(\lambda \mathbf{a}) = \lambda^n \det(\mathbf{a})$.

特别的，取 $\lambda=0$，我们得到

$$\det(\mathbf{0}, \mathbf{A}_2, \cdots, \mathbf{A}_n) = 0. \tag{67b}$$

当然，这同样适用于其他的列，因而我们得到，**如果方阵 a 有一列为零向量，它的行列式就等于零.** 从行列式的多线性性质，我们更一般地推得

$$\begin{aligned}
&\det(\mathbf{A}_1 + \lambda\mathbf{A}_2, \mathbf{A}_2, \cdots, \mathbf{A}_n) \\
=&\det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n) + \lambda\det(\mathbf{A}_2, \mathbf{A}_2, \cdots, \mathbf{A}_n) \\
=&\det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n),
\end{aligned} \tag{67c}$$

这是因为方阵 $(\mathbf{A}_2, \mathbf{A}_2, \cdots, \mathbf{A}_n)$ 有两个全同的列. 一般地，**当我们将方阵 a 的某一列的一个倍数加到另外一列上时，行列式的值不变**[1).]

具有根本的重要性的是行列式的乘法规则：

两个 n 阶方阵 a 与 b 的乘积的行列式等于它们的行列式的乘积：

$$\det(\mathbf{ab}) = \det(\mathbf{a}) \cdot \det(\mathbf{b}). \tag{68a}$$

用元素写出，这条规则具有如下形式：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \times \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix}, \tag{68b}$$

---

1) 显然，用因子 λ 乘 a 的某一列并加在该列上，相当于把行列式的值乘上 $(1+\lambda)$ 倍.

其中

$$c_{jk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}a_{nk} = \sum_{r=1}^{n} a_{jr}b_{rk}. \tag{68c}$$

这个规则是行列式定义的一个简单推论. 设 $\mathbf{c} = \mathbf{ab}$ 是乘积方阵. 我们让方阵 $\mathbf{a}$ 保持固定而考虑 $\mathbf{c}$ 的行列式对 $\mathbf{b}$ 的依赖性. 由 (68c) 方阵 $\mathbf{c}$ 的第 k 个列向量

$$\mathbf{C}_k = (c_{1k}, c_{2k}, \cdots, c_{nk})$$

有元素 c_{jk}, 它是方阵 $\mathbf{b}$ 的第 k 个列向量 $\mathbf{B}_k$ 的线性型. 由此推知, 当方阵 $\mathbf{b}$ 的其他列保持固定时, $\det(\mathbf{c})$ 是向量 $\mathbf{B}_k$ 的一个线性型. 另外, 交换 $\mathbf{b}$ 的两列正式好对应着交换 $\mathbf{c}$ 的相应列. 因此, $\det(\mathbf{c})$ 是方阵 $\mathbf{b}$ 的列向量的一个多线性交替型. 因之 (见第 184 页)

$$\det(\mathbf{c}) = \gamma \det(\mathbf{b}),$$

其中 γ 是 $\det(\mathbf{c})$ 的一个特殊值, 相当于特殊情形

$$\mathbf{B}_1 = \mathbf{E}_1,\ \mathbf{B}_2 = \mathbf{E}_2, \cdots, \mathbf{B}_n = \mathbf{E}_n,$$

即 $\mathbf{b}$ 是单位方阵 $\mathbf{e}$ 的情形. 现在, 取 $\mathbf{b} = \mathbf{e}$, 则显然有

$$\mathbf{c} = \mathbf{ab} = \mathbf{ae} = \mathbf{a},$$

因此 $\gamma = \det(\mathbf{a})$. 这就证明了 (68a).

在第 170 页, 我们把通过交换矩阵 $\mathbf{a}$ 的行和列得到的矩阵定义作 $\mathbf{a}$ 的转置矩阵 $\mathbf{a}^T$. 我们有一个出乎意外的事实, 即一个方阵与它的转置方阵具有相等的行列式

$$\det(\mathbf{a}^T) = \det(\mathbf{a}) \tag{68d}$$

或

$$\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{68e}$$

对于 $n = 2, 3$, 从第 175 页展开式 (51a), (52a), 很容易验证这个恒等式. 对于一般的 n, 我们仅仅提示一下证明, 它能够建立在 $\det(\mathbf{a})$ 的展开式 (66a) 的基础上. 对于和式中具有非零系数的每一项, 我们能按照第一个下标重排因子, 使得

$$a_{j_1 1} a_{j_2 2} \cdots a_{j_n n} = a_{1k_1} a_{2k_2} \cdots a_{nk_n},$$

其中 $k_1, k_2, \cdots, k_n$ 仍形成数字 $1, 2, \cdots, n$ 的一个置换[1]. 容易证明

$$\varepsilon_{j_1 j_2 \cdots j_n} = \varepsilon_{k_1 k_2 \cdots k_n}$$

(这个证明留给读者作为练习). 因此

$$\det(\mathbf{a}) = \sum_{k_1 \cdots k_n = 1}^{n} \varepsilon_{k_1 k_2 \cdots k_n} a_{1k_1} a_{2k_2} \cdots a_{nk_n} = \det(\mathbf{a}^T).$$

公式 (68d) 的一个直接推论是, 每个行列式能够看作它的行向量的一个多线性交替函数. 特别, **当交换行列式的任意两行时, 行列式变号.**

乘法规则 (68a) 说的是, **两个方阵 $\mathbf{a}, \mathbf{b}$ 的行列式的乘积等于方阵 $\mathbf{ab}$ 的行列式, 而方阵 $\mathbf{ab}$ 的元素就是 $\mathbf{a}$ 的行向量与 $\mathbf{b}$ 的列向量的数量积.** 现在, 我们利用这一事实, 即每个方阵 $\mathbf{a}$ 的行列式等于它的行列互换后所得到的转置方阵 $\mathbf{a}^T$ 的行列式. 由此推出

$$\det(\mathbf{a}) \cdot \det(\mathbf{b}) = \det(\mathbf{a}^T) \cdot \det(\mathbf{b}) = \det(\mathbf{a}^T \mathbf{b}).$$

因此, 方阵 $\mathbf{a}, \mathbf{b}$ **的行列式的乘积也等于方阵 $\mathbf{a}^T\mathbf{b}$ 的行列式, 而 $\mathbf{a}^T\mathbf{b}$ 的元素就是 $\mathbf{a}$ 的列向量与 $\mathbf{b}$ 的列向量的数量积.** 如果

$$\mathbf{a} = (\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n) \text{ 与 } \mathbf{b} = (\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_n),$$

---

1) 将 $j_1, j_2, \cdots, j_n$ 当作把集合 $1, 2, \cdots, n$ 映射到它自身上的一个函数, 就得到 $k_1, k_2, \cdots, k_n$ 正好是它的反函数, 就是说, 方程 $j_r = s$ 等价于 $k_s = r$.

我们就得到恒等式

$$\det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n) \cdot \det(\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_n) = \begin{vmatrix} \mathbf{A}_1 \cdot \mathbf{B}_1 & \mathbf{A}_1 \cdot \mathbf{B}_2 & \cdots & \mathbf{A}_1 \cdot \mathbf{B}_n \\ \mathbf{A}_2 \cdot \mathbf{B}_1 & \mathbf{A}_2 \cdot \mathbf{B}_2 & \cdots & \mathbf{A}_2 \cdot \mathbf{B}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_n \cdot \mathbf{B}_1 & \mathbf{A}_n \cdot \mathbf{B}_2 & \cdots & \mathbf{A}_n \cdot \mathbf{B}_n \end{vmatrix}. \tag{68f}$$

将这个规则直接应用于一个正交方阵 $\mathbf{a}$, 由于 [看第 170 页公式 (49)] 有 $\mathbf{a}^{-1} = \mathbf{a}^T$ 或 $\mathbf{a}\mathbf{a}^T = \mathbf{e}$, 便立即得到

$$\det(\mathbf{a}^T\mathbf{a}) = \det(\mathbf{a}^T)\det(\mathbf{a}) = [\det(\mathbf{a})]^2$$
$$= \det(\mathbf{e}) = 1.$$

因此, 正交方阵的行列式只取 $+1$ 或 -1 两个值. 这个结果的几何解释将在第 218 页给出.

e. 行列式对线性方程组的应用

行列式提供了一个方便的工具, 用以判定 n 维空间中 n 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 何时是相关的, 这也等价于判定以 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 为列向量的方阵何时是奇异的.

一个方阵是奇异的必要充分条件是它的行列式等于零.

先设 $\mathbf{a}$ 为一个奇异方阵. 于是列向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是相关的. 因此, 有一个列向量, 譬如说 $\mathbf{A}_1$, 相关于其他向量:

$$\mathbf{A}_1 = \lambda_2\mathbf{A}_2 + \lambda_3\mathbf{A}_3 + \cdots + \lambda_n\mathbf{A}_n.$$

由行列式的多线性性, 便有

$$\begin{aligned} \det(\mathbf{a}) &= \det(\lambda_2\mathbf{A}_2 + \lambda_3\mathbf{A}_3 + \cdots + \lambda_n\mathbf{A}_n, \mathbf{A}_2, \mathbf{A}_3, \cdots, \mathbf{A}_n) \\ &= \lambda_2\det(\mathbf{A}_2, \mathbf{A}_2, \cdots, \mathbf{A}_n) + \lambda_3\det(\mathbf{A}_3, \mathbf{A}_2, \mathbf{A}_3, \cdots, \mathbf{A}_n) \\ &\quad + \cdots + \lambda_n\det(\mathbf{A}_n, \mathbf{A}_2, \cdots, \mathbf{A}_n) = 0, \end{aligned}$$

因为这些方阵的每一个都有一个重复列[1]。

反之，如果 $\mathbf{a}$ 是非奇异的，则存在 (看第 168 页) $\mathbf{a}$ 的一个逆矩阵 $\mathbf{b}=\mathbf{a}^{-1}$ 使

$$\mathbf{ab}=\mathbf{e},$$

其中 $\mathbf{e}$ 为单位方阵. 由行列式的乘法规则得到

$$\det(\mathbf{a})\cdot\det(\mathbf{b})=\det(\mathbf{e})=1,$$

因此 $\det(\mathbf{a})\neq 0$. 这就证明了方阵 $\mathbf{a}$ 是奇异的充分必要条件为 $\det(\mathbf{a})=0$.

现在我们考虑对应于方阵 $\mathbf{a}$ 的线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=y_1, \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=y_2, \\ \qquad\cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=y_n. \end{cases} \tag{69a}$$

按照第 150 页的讨论，我们必须区分两种情形：

$$(1)\ \det(\mathbf{a})\neq 0, \qquad (2)\ \det(\mathbf{a})=0.$$

在情形 (1) 中，方程组 (69a) 对于每一组 $y_1,y_2,\cdots,y_n$ 有唯一的一组解. 在情形 (2) 中，不总是存在一组解，而且即使有解也决不是唯一的. 现在借助于行列式，我们不仅有一种明显的检验法来区分这两种情形，并且还将在情形 (1) 中求得计算解的方法. 引进向量

$$\mathbf{Y}=(y_1,y_2,\cdots,y_n),$$

我们能把方程组 (69a) 写成这个形式

$$x_1\mathbf{A}_1+x_2\mathbf{A}_2+\cdots+x_n\mathbf{A}_n=\mathbf{Y}, \tag{69b}$$

---

1) 更一般地，这个论证表明，当 $m>n$ 时，n 维空间中 m 个向量的多线性交替型恒等于零，因为这时这 m 个向量必定是相关的.

其中 $\mathbf{A}_k$ 为方阵 $\mathbf{a}$ 的列向量. 于是有

$$\begin{aligned}
&\det(\mathbf{Y},\mathbf{A}_2,\mathbf{A}_3,\cdots,\mathbf{A}_n)\\
=&\det(x_1\mathbf{A}_1+x_2\mathbf{A}_2+\cdots+x_n\mathbf{A}_n,\mathbf{A}_2,\mathbf{A}_3,\cdots,\mathbf{A}_n)\\
=&x_1\det(\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3,\cdots,\mathbf{A}_n)+x_2\det(\mathbf{A}_2,\mathbf{A}_2,\mathbf{A}_3,\cdots,\mathbf{A}_n)\\
&+x_3\det(\mathbf{A}_3,\mathbf{A}_2,\mathbf{A}_3,\cdots,\mathbf{A}_n)+\cdots+x_n\det(\mathbf{A}_n,\mathbf{A}_2,\mathbf{A}_3,\cdots,\mathbf{A}_n)\\
=&x_1\det(\mathbf{A}_1,\mathbf{A}_2,\mathbf{A}_3,\cdots,\mathbf{A}_n),
\end{aligned}$$

并且类似地有

$$\det(\mathbf{A}_1,\mathbf{Y},\mathbf{A}_3,\cdots,\mathbf{A}_n)=x_2\det(\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n),$$

等等. 如果方阵 $\mathbf{a}$ 是非奇异的, 我们就能用它的行列式去除, 并且得到解的行列式表示:

$$x_1=\frac{\det(\mathbf{Y},\mathbf{A}_2,\cdots,\mathbf{A}_n)}{\det(\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n)},\quad x_2=\frac{\det(\mathbf{A}_1,\mathbf{Y},\cdots,\mathbf{A}_n)}{\det(\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n)},\cdots,$$

$$x_n=\frac{\det(\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{Y})}{\det(\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n)}.$$

这就是关于解 n 个未知数、 n 个线性方程的方程组的**克拉默法则**.

练 习 2.3

1. 计算下面的行列式

(a) $\begin{vmatrix}3&4&5\\4&5&6\\5&6&7\end{vmatrix}$, (b) $\begin{vmatrix}1&1&1\\1&2&4\\1&3&9\end{vmatrix}$,

(c) $\begin{vmatrix}1&1&1\\2&3&4\\3&-1&7\end{vmatrix}$, (d) $\begin{vmatrix}1&x&x^3\\1&y&y^3\\1&z&z^3\end{vmatrix}$.

2. 求出 a,b,c 之间必定存在的关系，使得方程组

$$\begin{cases} 3x+4y+5z=a \\ 4x+5y+6z=b \\ 5x+6y+7z=c \end{cases}$$

可以有一组解.

3. (a) 验证单位方阵的行列式为 1.

(b) 证明，如果 $\mathbf{a}$ 是非奇异的方阵，则有

$$\det(\mathbf{a}^{-1}) = 1/\det(\mathbf{a}).$$

4. 求下列 ϵ 之值

(a) ε_{321}，(b) ε_{2143}，(c) ε_{4231}，(d) ε_{54321}.

5. 考虑这样两个手续：(1) 交换两行或两列；(2) 把某行 (或列) 的一个倍数加到另一行 (或列). 证明仅仅重复运用这两个手续便能把行列式

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix}$$

化成这个形式：

$$\begin{vmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{vmatrix}.$$

6. 如果方阵 $\mathbf{a}$ 的元素 $a_{ij}=0$, 只要 $i \neq j$, 则称 $\mathbf{a}$ 为一个对角方阵. 证明 $n \times n$ 对角方阵 (a_{ij}) 的行列式等于乘积 $a_{11}a_{22}\cdots a_{nn}$.

7. 方阵 (a_{ij}) 称为一个上三角方阵，如果 $j<i$ 时有 $a_{ij}=0$. 证明

$$\det(a_{ij}) = a_{11}a_{22}\cdots a_{nn}.$$

8. 计算

(a) $\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}$,

(b) $\begin{vmatrix} 1! & 2! & 3! \\ 2! & 3! & 4! \\ 3! & 4! & 5! \end{vmatrix}$,

(c) $\begin{vmatrix} 1! & 2! & 3! & 4! \\ 2! & 3! & 4! & 5! \\ 3! & 4! & 5! & 6! \\ 4! & 5! & 6! & 7! \end{vmatrix}$.

9. 解方程组

$$\begin{cases} 2x - 3y + 4z = 4, \\ 4x - 9y + 16z = 10, \\ 8x - 27y + 64z = 34. \end{cases}$$

10. 通过行列式

$$\begin{vmatrix} a & b \\ -b & a \end{vmatrix} \text{ 与 } \begin{vmatrix} c & d \\ -d & c \end{vmatrix}$$

相乘，证明恒等式

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (bc - ad)^2.$$

11. 设 $A = x^2 + y^2 + z^2$, $B = xy + yz + zx$, 证明

$$D = \begin{vmatrix} B & A & B \\ B & B & A \\ A & B & B \end{vmatrix} = (x^3 + y^3 + z^3 - 3xyz)^2.$$

12. 证明

$$\Delta = \begin{vmatrix} t_1 + x & a + x & a + x & a + x \\ b + x & t_2 + x & a + x & a + x \\ b + x & b + x & t_3 + x & a + x \\ b + x & b + x & b + x & t_4 + x \end{vmatrix}$$

具有形式 $A + Bx$, 其中 A, B 与 x 无关. 试给 x 以特定的值来证明

$$A = \frac{af(b) - bf(a)}{a - b}, \quad B = \frac{f(b) - f(a)}{b - a},$$

其中

$$f(t) = (t_1 - t)(t_2 - t)(t_3 - t)(t_4 - t).$$

13. 证明向量 $\mathbf{A}, \mathbf{B}$ 的任何一个双线性型 f 都可以写成

$$\mathbf{A}\cdot(\mathbf{cB})=(\mathbf{c}^T\mathbf{A})\cdot\mathbf{B}.$$

14. 证明，一个非奇异的仿射变换将每个二次曲面

$$ax^2+by^2+cz^2+dxy+exz+fyz+gx+hy+iz+j=0$$

映到另一个二次曲面.

15. 如果这三个行列式

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}, \quad \begin{vmatrix} a_1 & a_2 \\ c_1 & c_2 \end{vmatrix}, \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}$$

不全为零，则方程组

$$a_1x+a_2y=d,$$

$$b_1x+b_2y=e,$$

$$c_1x+c_2y=f$$

存在一组解的必要充分条件是

$$D=\begin{vmatrix} a_1 & a_2 & d \\ b_1 & b_2 & e \\ c_1 & c_2 & f \end{vmatrix}=0.$$

16. 叙述两条直线

$$x=a_1t+b_1, \quad y=a_2t+b_2, \quad z=a_3t+b_3$$

与

$$x=c_1t+d_1, \quad y=c_2t+d_2, \quad z=c_3t+d_3$$

相交或平行的条件.

17. 不管展开式 (66a) 中每一项的因子是按照它们的第一个或第二个下标排序，即

$$a_{j_1 1}a_{j_2 2}\cdots a_{j_n n}=a_{1k_1}a_{2k_2}\cdots a_{nk_n},$$

总有

$$\varepsilon_{j_1 j_2 \cdots j_n} = \varepsilon_{k_1 k_2 \cdots k_n}.$$

由此证明 (68d).

18. 证明仿射变换

$$x' = ax + by + cz$$
$$y' = dx + ey + fz$$
$$z' = gx + hy + kz$$

保持至少一个方向不改变.

2.4　行列式的几何解释

a. 向量积与三维空间中平行六面体的体积

在第一卷第 431 页我们把两个平面向量 $\mathbf{A} = (a_1, a_2)$ 与 $\mathbf{B} = (b_1, b_2)$ 的交叉积定义为数量

$$\mathbf{A} \times \mathbf{B} = a_1 b_2 - a_2 b_1. \tag{70a}$$

这个数量的绝对值 $|\mathbf{A} \times \mathbf{B}|$ 是以 P_0, P_1, P_2 为顶点的三角形的面积的两倍，其中 $\mathbf{A} = \overrightarrow{P_0P_1}$, $\mathbf{B} = \overrightarrow{P_0P_2}$. 我们称 $|\mathbf{A} \times \mathbf{B}|$ 为由向量 $\mathbf{A}$ 与 $\mathbf{B}$ 所**张成**的平行四边形的面积. 这个平行四边形有相继顶点 P_0, P_1, Q, P_2. $\mathbf{A} \times \mathbf{B}$ 的符号决定了平行四边形的定向 [1]. 引用行列式的记号，交叉积有这样的形式:

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \det(\mathbf{A}, \mathbf{B}). \tag{70b}$$

因此，$|\det(\mathbf{A}, \mathbf{B})|$ 能几何地解释为由向量 $\mathbf{A}, \mathbf{B}$ 所张成的平行四边形的面积. 我们将建立关于高阶行列式的类似的解释.

---

1) 如果顶点顺次序的定向 (反时针或顺时针) 与 "坐标正方形" 的相继顶点 $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$ 的定向是一样的，就有 $\mathbf{A} \times \mathbf{B} > 0$.

对于三维空间的三个向量 $\mathbf{A}=(a_1,a_2,a_3)$, $\mathbf{B}=(b_1,b_2,b_3)$, $\mathbf{C}=(c_1,c_2,c_3)$, 可自然地作三阶行列式

$$\det(\mathbf{A},\mathbf{B},\mathbf{C})=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

写成向量 $\mathbf{C}$ 的线性型, 从 (52a) 我们有

$$\begin{aligned}\det(\mathbf{A},\mathbf{B},\mathbf{C}) &= (a_2b_3-a_3b_2)c_1+(a_3b_1-a_1b_3)c_2+(a_1b_2-a_2b_1)c_3 \\ &= \mathbf{Z}\cdot\mathbf{C},\end{aligned} \tag{71a}$$

其中 $\mathbf{Z}=(z_1,z_2,z_3)$ 是具有下列分量的向量:

$$\begin{aligned} z_1 &= a_2b_3-a_3b_2=\begin{vmatrix} a_2 & b_2 \\ a_3 & b_3\end{vmatrix}, \\ z_2 &= a_3b_1-a_1b_3=\begin{vmatrix} a_3 & b_3 \\ a_1 & b_1\end{vmatrix}, \\ z_3 &= a_1b_2-a_2b_1=\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2\end{vmatrix}. \end{aligned} \tag{71b}$$

我们称向量 $\mathbf{Z}$ 是向量 $\mathbf{A},\mathbf{B}$ 的 "向量积" 或 "交叉积", 并写作

$$\mathbf{Z}=\mathbf{A}\times\mathbf{B}\ ^{1)}.$$

于是由定义有

$$\det(\mathbf{A},\mathbf{B},\mathbf{C})=(\mathbf{A}\times\mathbf{B})\cdot\mathbf{C}. \tag{71c}$$

由于有这个公式, 有时把数量 $\det(\mathbf{A},\mathbf{B},\mathbf{C})$ 叫作 $\mathbf{A},\mathbf{B},\mathbf{C}$ 的 **三重向量积**.

向量 $\mathbf{Z}=\mathbf{A}\times\mathbf{B}$ 的分量 z_i 本身就是二阶行列式, 因而是向量 $\mathbf{A},\mathbf{B}$ 的双线性交替型. 由此立即推出关于向量乘法的规则

$$(\lambda\mathbf{A}\times\mathbf{B})=\mathbf{A}\times(\lambda\mathbf{B})=\lambda(\mathbf{A}\times\mathbf{B}), \tag{72a}$$

---

1) 两个三维向量的交叉积仍是一个 **向量**, 这与二维向量的交叉积不同, 也与任何维数向量的数量积不同 (都是 **数量**).

$$(\mathbf{A}' + \mathbf{A}'') \times \mathbf{B} = \mathbf{A}' \times \mathbf{B} + \mathbf{A}'' \times \mathbf{B}, \tag{72b}$$
$$\mathbf{A} \times (\mathbf{B}' + \mathbf{B}'') = \mathbf{A} \times \mathbf{B}' + \mathbf{A} \times \mathbf{B}'',$$
$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}. \tag{72c}$$

关系式 (72c) 可称为乘法的"负交换律". 它有重要的推论:

$$\mathbf{A} \times \mathbf{A} = 0, \text{ 对于所有的 } \mathbf{A}. \tag{72d}$$

更一般地, **向量积 $\mathbf{A} \times \mathbf{B}$ 等于零当且仅当 $\mathbf{A}$ 与 $\mathbf{B}$ 是相关的**. 因为由 (71c), 关系式 $\mathbf{A} \times \mathbf{B} = 0$ 等价于

$$\det(\mathbf{A}, \mathbf{B}, \mathbf{C}) = 0 \text{ 对于所有的 } \mathbf{C},$$

或者说 (看第 189 页), 对于所有的 $\mathbf{C}$ 而言, $\mathbf{A}, \mathbf{B}, \mathbf{C}$ 都是相关的. 然而, 我们总能找到一个向量 $\mathbf{C}$, 它对于 $\mathbf{A}, \mathbf{B}$ 是无关的 (看第 150 页); 所以 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ 的相关性隐含着 $\mathbf{A}$ 与 $\mathbf{B}$ 是相关的.

向量积 $\mathbf{A} \times \mathbf{B}$ 是垂直于 $\mathbf{A}$ 与 $\mathbf{B}$ 的, 因为由 (71c) 有

$$\begin{aligned} (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{A} &= \det(\mathbf{A}, \mathbf{B}, \mathbf{A}) = 0, \\ (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{B} &= \det(\mathbf{A}, \mathbf{B}, \mathbf{B}) = 0. \end{aligned} \tag{72e}$$

因此, 对于无关的向量 $\mathbf{A} = \overrightarrow{P_0P_1}$ 与 $\mathbf{B} = \overrightarrow{P_0P_2}$, 它们所张成的平面 $P_0P_1P_2$ 有两个垂直方向, $\mathbf{A} \times \mathbf{B}$ 是其中之一. 向量 $\mathbf{A} \times \mathbf{B}$ 的长度同样有一个简单的几何解释. 由 (71b), 我们有

$$\begin{aligned} |\mathbf{A} \times \mathbf{B}|^2 &= (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2 \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\ &= |\mathbf{A}|^2|\mathbf{B}|^2 - (\mathbf{A} \cdot \mathbf{B})^2 \, {}^{1)}. \end{aligned} \tag{72f}$$

利用

$$\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}||\mathbf{B}| \cos \gamma$$

---

1) 从这个恒等式可附带地直接证明柯西 - 施瓦茨不等式 $|\mathbf{A} \cdot \mathbf{B}| \le |\mathbf{A}||\mathbf{B}|$ (看第 143 页). 它还提供另外一项知识, 就是当且仅当 $\mathbf{A}, \mathbf{B}$ 相关时等号成立.

这个事实, 其中 γ 是 $\mathbf{A}$ 与 $\mathbf{B}$ 的方向之间的夹角 (看第 141 页公式 (14)), 我们由 (72f) 得到

$$|\mathbf{A}\times\mathbf{B}| = \sqrt{|\mathbf{A}|^2|\mathbf{B}|^2 - |\mathbf{A}|^2|\mathbf{B}|^2\cos^2\gamma} = |\mathbf{A}||\mathbf{B}|\sin\gamma.$$

对于 $\mathbf{A} = \overrightarrow{P_0P_1}$, $\mathbf{B} = \overrightarrow{P_0P_2}$, 我们有从直线 P_0P_1 到点 P_2 的距离 (图 2.6) $|\mathbf{B}|\sin\gamma$ (γ 取 0 和 π 之间的一个值). 因此 (与二维情形正好一样), 量 $|\mathbf{A}\times\mathbf{B}|$ 给出由 $\mathbf{A}$ 与 $\mathbf{B}$ 所 "张成" 具有顶点 P_0, P_1, Q, P_2 的平行四边形的面积, 或说是以 P_0, P_1, P_2 为顶点的三角形面积的两倍.

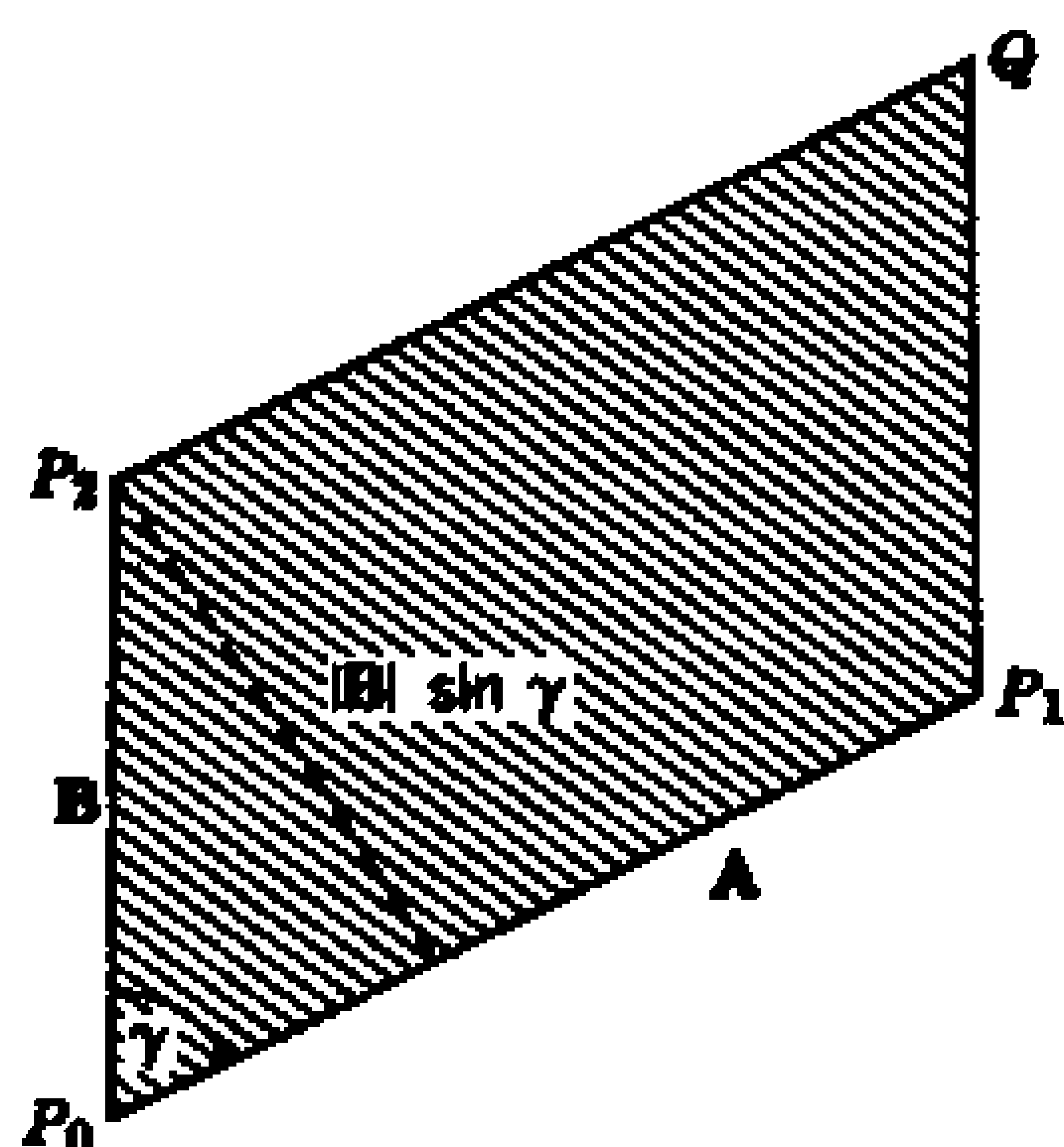


图 2.6　两个向量 $\mathbf{A}, \mathbf{B}$ 所张的平行四边形的面积 $|\mathbf{A}\times\mathbf{B}|$

向量积 $\mathbf{A}\times\mathbf{B} = (z_1, z_2, z_3)$ 的每个分量也能够从几何上解释. 例如, 表达式

$$z_3 = a_1b_2 - a_2b_1$$

恰好是二维向量 (a_1, a_2) 与 (b_1, b_2) 的交叉积 [看 (70a)]. 如果 P_0 有坐标 ξ_1, ξ_2, ξ_3, 则在 (x_1, x_2) 平面上具有顶点 $(\xi_1, \xi_2), (\xi_1 + a_1, \xi_2 + a_2), (\xi_1 + a_1 + b_1, \xi_2 + a_2 + b_2), (\xi_1 + b_1, \xi_2 + b_2)$ 的平行四边形的面积就是 $|z_3|$. 这个平行四边形恰好是向量 $\mathbf{A}, \mathbf{B}$ 在空间中所张成的以 P_0, P_1, Q, P_2 为顶点的平行四边形在 (x_1, x_2) 平面上的投影 (看图 2.7). 如果 $\mathbf{A}\times\mathbf{B}$ 有方向余弦 $\cos\beta_1, \cos\beta_2, \cos\beta_3$, 我们有 (看第

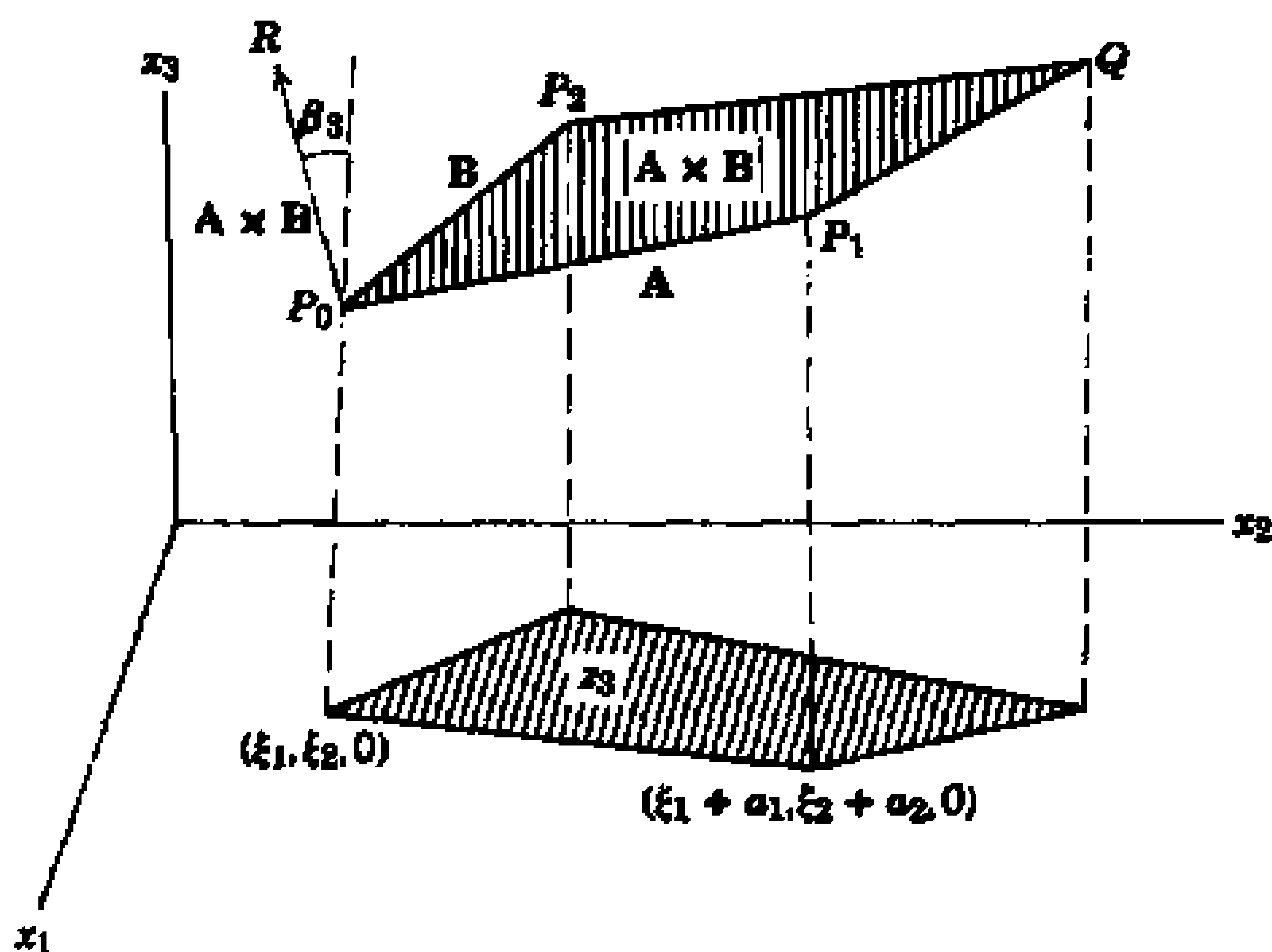


图 2.7　向量积 $\mathbf{A}\times\mathbf{B}=(z_1,z_2,z_3)$ 的分量解释为投影面积

139 页 (9))

$$|z_3| = |\mathbf{A}\times\mathbf{B}||\cos\beta_3|.$$

因此，$\cos\beta_3$ 给出了 $\mathbf{A},\mathbf{B}$ 所张成的平行四边形在 (x_1,x_2) 平面投影的面积相对于它自身的面积之比. 这里 β_3 是 x_3 轴与经过 P_0,P_1,P_2 的平面的法线之间的夹角. 当然，这个角度与包含 $\mathbf{A},\mathbf{B}$ 所张成的平行四边形的平面与 (x_1,x_2) 平面之间的夹角是相同的 [1].

如果 $\mathbf{A}=\overrightarrow{P_0P_1}$ 与 $\mathbf{B}=\overrightarrow{P_0P_2}$ 是无关的向量，我们有

$$\mathbf{A}\times\mathbf{B}=\overrightarrow{P_0R},$$

其中点 R 位于经过 P_0 而垂直于平面 $P_0P_1P_2$ 的直线上，并且与 P_0 的距离是三角形 $P_0P_1P_2$ 面积的两倍. 这几乎唯一地确定了 R. 具有这样性质的点只有两个，位于该平面的两边. 这两个点中究竟哪一个是向量 $\mathbf{A}\times\mathbf{B}=\overrightarrow{P_0R}$ 的末点 R, 能用下面的"连续性"论点来确定. 因为向量积 $\mathbf{A}\times\mathbf{B}$ 的分量是向量 $\mathbf{A},\mathbf{B}$ 的双线性函数，所以向量积 $\mathbf{A}\times\mathbf{B}$ 连续地依赖于 $\mathbf{A},\mathbf{B}$. 于是，只要

$$\mathbf{A}\times\mathbf{B}\neq 0,$$

1) 一般地，一个平面图形在第二个平面上的投影的面积等于原来图形的面积与这两个平面之间夹角的余弦之积. 当我们讨论积分变换时，就会明白这一点.

就是说只要 $\mathbf{A}$ 与 $\mathbf{B}$ 保持不变成 $\mathbf{O}$ 或平行，则 $\mathbf{A}\times\mathbf{B}$ 的方向也就连续地依赖于 $\mathbf{A},\mathbf{B}$. 我们总能按照如下方式连续地变动这两个向量 $\mathbf{A}$ 与 $\mathbf{B}$, 使它们决不为 $\mathbf{O}$ 或平行，直到最后 $\mathbf{A}$ 与坐标向量 $\mathbf{E}_1=(1,0,0)$ 重合， $\mathbf{B}$ 与坐标向量 $\mathbf{E}_2=(0,1,0)$ 重合. 这等于把三角形 $P_0P_1P_2$ 连续地非退化地变形，使得 P_0 变到原点，而 P_1 与 P_2 分别变到正 x_1 轴与正 x_2 轴上原点距离为 1 的点. 在这个过程中，经过 P_0 而垂直于平面 $P_0P_1P_2$ 的直线上的点 R 并不穿过该平面. 现在，由 (71b) 有

$$\mathbf{E}_1\times\mathbf{E}_2=(0,0,1)=\mathbf{E}_3.$$

在常用的"右手"坐标系中， $\mathbf{E}_3$ 是垂直于 $\mathbf{E}_1$ 和 $\mathbf{E}_2$ 并且按照如下方式来确定的，即从 $(0,0,1)$ 看去， $\mathbf{E}_1$ 关于 x_3 轴逆时针旋转 $90°$ 就到达 $\mathbf{E}_2$. 于是，一般地，如果我们的坐标系是右手系，

$$\mathbf{A}\times\mathbf{B}=\overrightarrow{P_0R}$$

的方向就是这样确定的，即从 R 看时，向量 $\mathbf{A}=\overrightarrow{P_0P_1}$ 关于直线 $\overrightarrow{P_0R}$ 逆时针旋转一个位于 0 与 π 之间的角度 γ 就到达 $\mathbf{B}=\overrightarrow{P_0P_2}$ (图 2.9). 类似地，在左手坐标系中，从 (0,0,1) 看去 $\mathbf{E}_1$ 是顺时针旋转 $90°$ 到达 $\mathbf{E}_2$, 同样从 $\mathbf{A}\times\mathbf{B}=\overrightarrow{P_0R}$ 的末点 R 看去，从 $\mathbf{A}$ 到 $\mathbf{B}$ 也是这样旋转.

一般地，三个无关向量 $\mathbf{A},\mathbf{B},\mathbf{C}$ 的有序组定义一个确定的**指向**或**定向**. 如果

$$\mathbf{A}=\overrightarrow{P_0P_1},\quad \mathbf{B}=\overrightarrow{P_0P_2},\quad \mathbf{C}=\overrightarrow{P_0P_3},$$

我们总能在平面 $P_0P_1P_2$ 内将 $\mathbf{A}$ 的方向转动一个 0 到 π 之间的角度而达到 $\mathbf{B}$ 的方向. 从向量 $\mathbf{C}$ 指向的平面 $P_0P_1P_2$ 的一侧来看，旋转所具有的定向就定义作三重组 $\mathbf{A},\mathbf{B},\mathbf{C}$ 的定向 (顺时针或逆时针)[1]. 三重组 $\mathbf{B},\mathbf{A},\mathbf{C}$ 有**相反的**定向. **三重组** $\mathbf{A},\mathbf{B},\mathbf{A}\times\mathbf{B}$ **与坐**

1) 这同一类型定向的方式也确定着左手螺旋与右手螺旋之间的区别. 一个螺旋运动是由沿着一个轴的平移与关于这个轴的旋转所组成的. 这两种类型螺旋的区别取决于，当从轴的平移进行的方向看去时，该旋转所具有的定向 (顺时针或逆时针).

标向量 $\mathbf{E}_1,\mathbf{E}_2,\mathbf{E}_3$ 的方向永远是相同的.

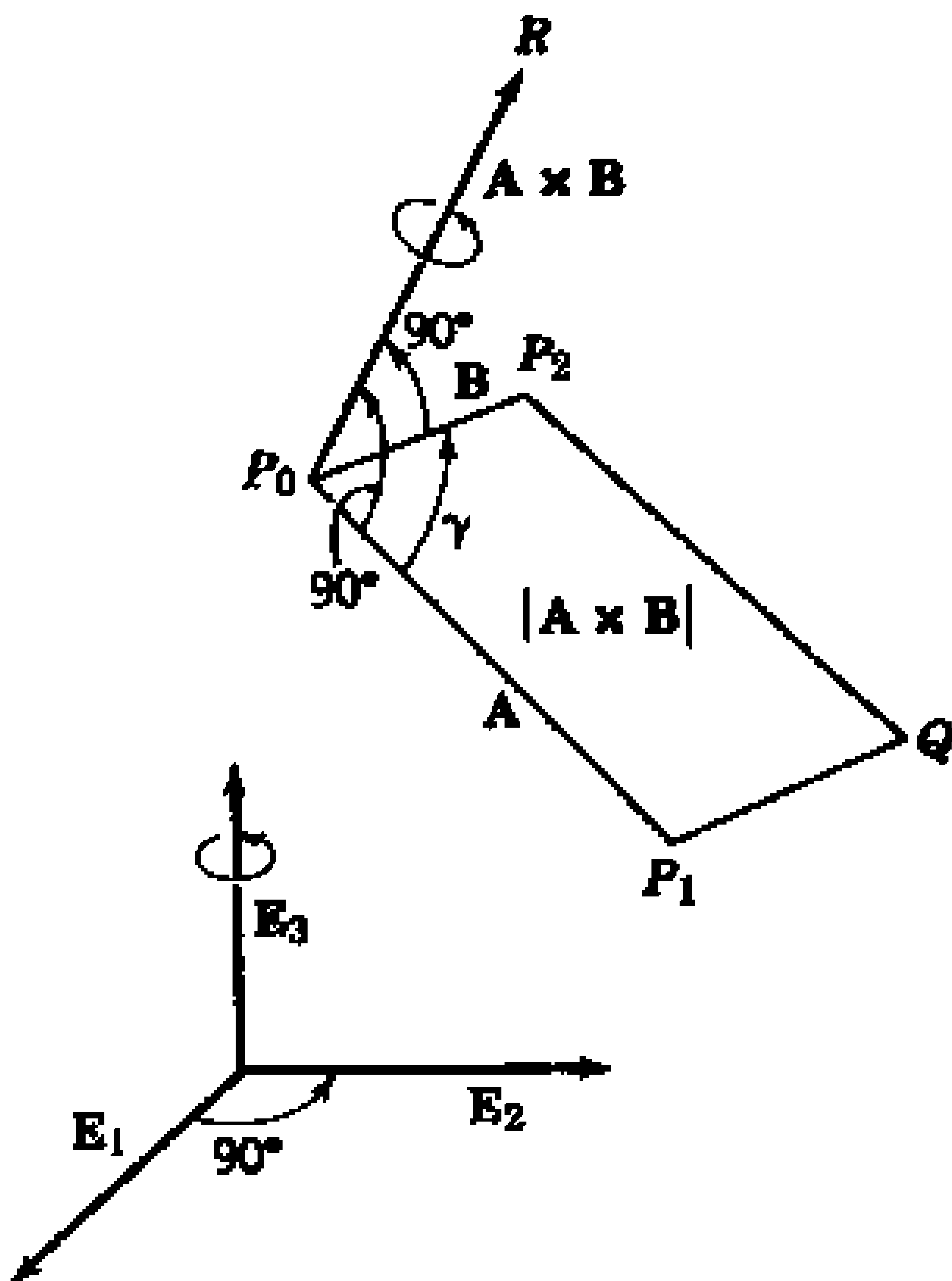


图 2.8　右手坐标系中的向量积 $\mathbf{A}\times\mathbf{B}$

如果三重组 $\mathbf{A},\mathbf{B},\mathbf{C}$ 与坐标向量的三重组 $\mathbf{E}_1,\mathbf{E}_2,\mathbf{E}_3$ 有同样的定向, 我们就称三重组 $\mathbf{A},\mathbf{B},\mathbf{C}$ 相对于 (x_1,x_2,x_3) 坐标系有正向; 如果有相反的方向, 就称之为负向. 三重组 $\mathbf{A},\mathbf{B},\mathbf{C}$ 有正向的必要充分条件是

$$\det(\mathbf{A},\mathbf{B},\mathbf{C})>0. \tag{73}$$

因为关系式 (73) 意味着

$$(\mathbf{A}\times\mathbf{B})\cdot\mathbf{C}>0,$$

也就是说, 向量 $\mathbf{A}\times\mathbf{B}$ 的方向与向量 $\mathbf{C}$ 的方向作成一个锐角. 若令

$$\mathbf{A}=\overrightarrow{P_0P_1},\quad \mathbf{B}=\overrightarrow{P_0P_2},\quad \mathbf{C}=\overrightarrow{P_0P_3},$$

则 $\mathbf{A}\times\mathbf{B}$ 垂直于平面 $P_0P_1P_2$, 而这隐含着向量 $\mathbf{A}\times\mathbf{B}$ 与向量 $\mathbf{C}$ 指向平面的同一侧. 因此, $\mathbf{A},\mathbf{B},\mathbf{C}$ 与 $\mathbf{A},\mathbf{B},\mathbf{A}\times\mathbf{B}$ 有相同的定向, 这就是 $\mathbf{E}_1,\mathbf{E}_2,\mathbf{E}_3$ 的定向.

三个无关的向量 $\mathbf{A},\mathbf{B},\mathbf{C}$, 当给定同一个起点 P_0 时, 它们 "张成" 一个确定的平行六面体, 它以 $\mathbf{A},\mathbf{B},\mathbf{C}$ 的末点为相邻于 P_0 的顶点. 按照三重组 $\mathbf{A},\mathbf{B},\mathbf{C}$ 的定向, 我们称平行六面体对于 (x_1,x_2,x_3) 坐标系而言有正向或负向. 交换向量 $\mathbf{A},\mathbf{B},\mathbf{C}$ 中的任意两个就改变了它们所张成的平行六面体的定向 [1].

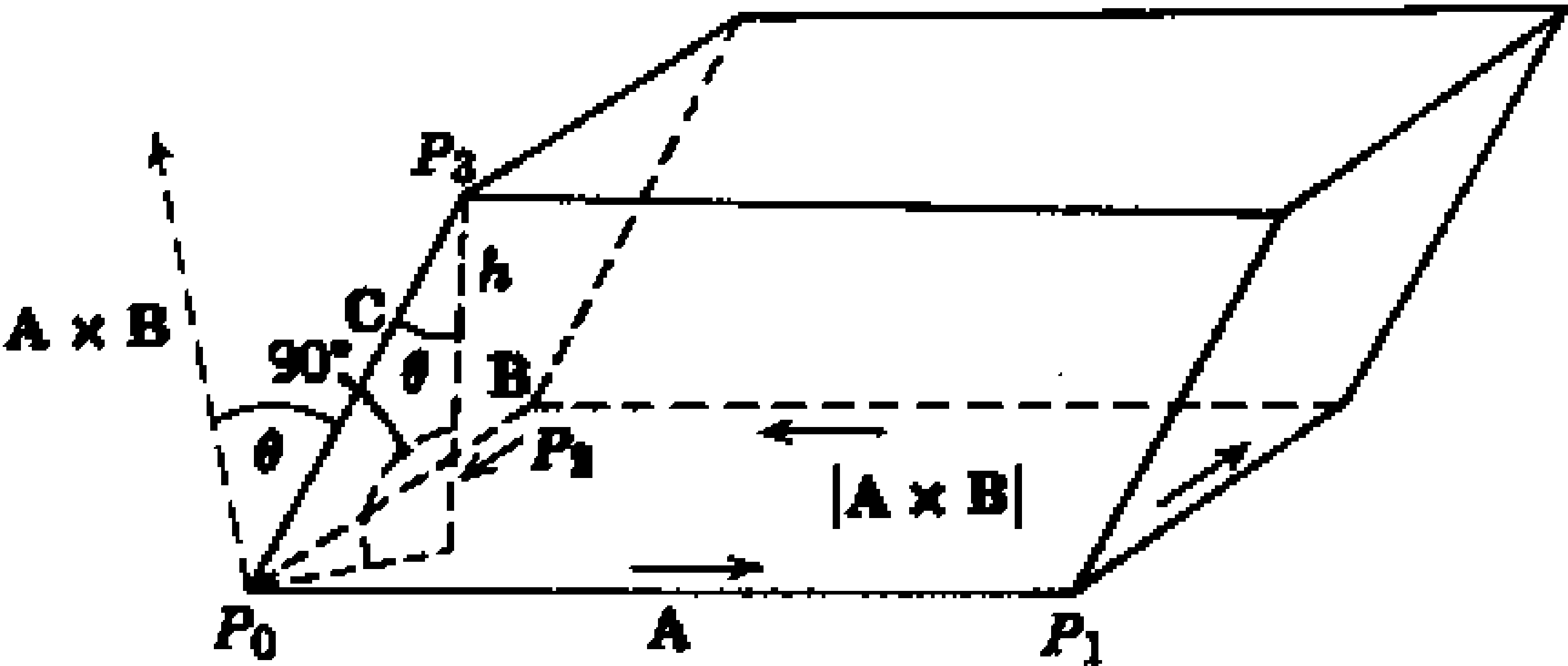


图 2.9 平行六面体的体积 $V=|\mathbf{A}\times\mathbf{B}|h$

设 θ 是向量 $\mathbf{C}$ 与 $\mathbf{A}\times\mathbf{B}$ 的方向所构成的角度. 由 (71c) 有

$$\det(\mathbf{A},\mathbf{B},\mathbf{C})=|\mathbf{A}\times\mathbf{B}||\mathbf{C}|\cos\theta. \tag{74a}$$

由于 $\mathbf{A}\times\mathbf{B}$ 垂直于平面 $P_0P_1P_2$, 直线 P_0P_3 与平面 $P_0P_1P_2$ 之间的夹角是 $\frac{\pi}{2}-\theta$. 因此, 点 P_3 到平面 $P_0P_1P_2$ 的距离是

$$h=|\mathbf{C}||\cos\theta|=|\mathbf{C}|\left|\sin\left(\frac{\pi}{2}-\theta\right)\right|. \tag{74b}$$

这是平行六面体 P_3-$P_0P_1P_2$ 的**高度**. 因为平行六面体的体积 V 等于它的一个面的面积 $|\mathbf{A}\times\mathbf{B}|$ 乘上对应的高 h, 我们从 (74a, b) 得到

$$V=|\mathbf{A}\times\mathbf{B}|h=|\det(\mathbf{A},\mathbf{B},\mathbf{C})|. \tag{74c}$$

用话来说, **三个向量** $\mathbf{A},\mathbf{B},\mathbf{C}$ **所张成的平行六面体的体积等于以** $\mathbf{A},\mathbf{B},\mathbf{C}$ **为列向量之方阵的行列式的绝对值**. 因此, $\det(\mathbf{A},\mathbf{B},\mathbf{C})$

---

1) 平行六面体的定向可具体化为赋予平行六面体的每个面一种定向 (即规定每个面的边界多边形的一个指向), 使得两个相邻的面的公共棱按照两个面上的定向而具有相反的指向. 如果对于单独一个面的一个棱确定了指向, 则各面的定向都被唯一地确定了. 对于 $\mathbf{A},\mathbf{B},\mathbf{C}$ 所张成的平行六面体的定向来说, P_0P_1 作为 $\overrightarrow{P_0P_2}$ 与 $\overrightarrow{P_0P_1}$ 所张的面的一个棱其指向是从 P_0 到 P_1 (看图 2.9).

的值同时确定了由 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ 所张成的平行六面体的体积及其定向. 我们用如下公式来表达这个事实

$$\det(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \varepsilon V, \tag{74}$$

其中 V 是由 $\mathbf{A}, \mathbf{B}, \mathbf{C}$ 所张成的平行六面体的体积，并且当这个平行六面体相对于 (x_1, x_2, x_3) 坐标系有正的定向时 $\varepsilon = +1$, 而有负的定向时 $\varepsilon = -1$.

b. 行列式关于一列的展开式. 高维向量积

仅仅在三维空间中, 我们能定义两个向量 $\mathbf{A}, \mathbf{B}$ 的一个乘积 $\mathbf{A} \times \mathbf{B}$ 仍为一个向量 [1]. 在 n 维空间中最密切的模拟也许是 $n-1$ 个向量的 "向量积". 试在 n 维空间中取 n 个向量

$$\mathbf{A}_1 = (a_{11}, \cdots, a_{n1}), \cdots, \mathbf{A}_n = (a_{1n}, \cdots, a_{nn}).$$

我们能够作出以这些向量为列向量的方阵的行列式, 它是最后一个向量 $\mathbf{A}_n$ 的线性型，因而能写成一个数量积

$$\begin{aligned} \det(\mathbf{A}_1, \cdots, \mathbf{A}_n) &= z_1 a_{1n} + z_2 a_{2n} + \cdots + z_n a_{nn} \\ &= \mathbf{Z} \cdot \mathbf{A}_n, \end{aligned} \tag{75}$$

其中向量 $\mathbf{Z} = (z_1, z_2, \cdots, z_n)$ 仅仅依赖于前 $n-1$ 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{n-1}$. 显然， $\mathbf{Z}$ 关于 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{n-1}$ 中的每一个都分别是线性的并且是交替的. 我们可以称 $\mathbf{Z}$ 为 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{n-1}$ 的 **向量积** 并记作

$$\mathbf{Z} = \mathbf{A}_1 \times \mathbf{A}_2 \times \cdots \times \mathbf{A}_{n-1}. \tag{76}$$

由 (75) 清楚地看出:

$$\mathbf{Z} \cdot \mathbf{A}_1 = \mathbf{Z} \cdot \mathbf{A}_2 = \cdots = \mathbf{Z} \cdot \mathbf{A}_{n-1} = 0;$$

---

1) 在高维空间中我们不能以 **几何** 方式在向量 $\mathbf{A}, \mathbf{B}$ 所张的平面之外定义一个第三向量 $\mathbf{C}$ 来与 $\mathbf{A}, \mathbf{B}$ 联系，就是说，不能通过一个作图法由 $\mathbf{A}, \mathbf{B}$ 唯一地确定一个向量 $\mathbf{C}$, 使其在刚体运动下不变.

可见与三维的情形一样，$n-1$ 个向量的向量积与这 $n-1$ 个向量的每一个都是正交的. 我们不久将看到，这向量积的长度也可以几何地解释为向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_{n-1}$ 所张成的定向的 $n-1$ 维平行多面体的体积.

与三维的情形一样，$\mathbf{Z}$ 的分量能写成类似于公式 (71b) 的行列式. 让我们先对 $\mathbf{Z}$ 的分量 z_n 导出这样一个行列式的表达式. 由 (75) 有

$$z_n = \mathbf{Z} \cdot \mathbf{E}_n = \det(\mathbf{A}_1, \cdots, \mathbf{A}_{n-1}, \mathbf{E}_n),$$

其中

$$\mathbf{E}_n = (0, \cdots, 0, 1)$$

是第 n 个坐标向量. 在第 185 页关于行列式的一般展开公式 (66a) 中取 $\mathbf{A}_n = \mathbf{E}_n$, 这等于将每一项的最后一个因子 $a_{j_n n}$ 换成 1, 当 $j_n = n$ 时；换成 0, 当 $j_n \neq n$ 时. 对于 $j_n = n$, 系数 $\varepsilon_{j_1 j_2 \cdots j_n}$ 总等于 0, 除非 $j_1, j_2, \cdots, j_{n-1}$ 是 $1, 2, \cdots, n-1$ 的一个排列. 而在这种情形，系数 (65c, d) 就变成

$$\begin{aligned}\varepsilon_{j_1 \cdots j_{n-1} j_n} &= \varepsilon_{j_1 \cdots j_{n-1} n} = \operatorname{sgn} \phi(j_1, \cdots, j_{n-1}, n) \\ &= \operatorname{sgn}(n - j_{n-1}) \cdots (n - j_1) \phi(j_1, \cdots, j_{n-1}) \\ &= \operatorname{sgn} \phi(j_1, j_2, \cdots, j_{n-1}) = \varepsilon_{j_1 \cdots j_{n-1}}.\end{aligned}$$

于是由 (66a) 得到

$$\begin{aligned} z_n &= \sum_{j_1 \cdots j_{n-1}=1}^{n-1} \varepsilon_{j_1 j_2 \cdots j_{n-1}} a_{j_1 1} a_{j_2 2} \cdots a_{j_{n-1} n-1} \\ &= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1\,n-1} \\ a_{21} & a_{22} & \cdots & a_{2\,n-1} \\ \vdots & \vdots & & \vdots \\ a_{(n-1)\,1} & a_{(n-1)\,2} & \cdots & a_{(n-1)\,(n-1)} \end{vmatrix}. \end{aligned} \tag{77a}$$

可见，z_n 等于从方阵 $(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n)$ 中去掉最后一行最后一列所得到的行列式. 一般地，人们把从一个方阵 $\mathbf{a}$ 去掉它的一些行和

列，而保持剩下元素的相对位置，所得到的方阵的行列式定义作矩阵 $\mathbf{a}$ 的一个**子式**．一个方阵 $\mathbf{a}$ 的**一个元素** a_{jk} **的余子式**便是从 $\mathbf{a}$ 中去掉包含 a_{jk} 的行和列所得到的行列式．**这样**，z_n **就等于** a_{nn} **的余子式**．

向量 $\mathbf{Z}$ 的其他分量都有类似的表示式．例如，由 (75) 我们有

$$z_{n-1}=\det(\mathbf{A}_1,\cdots,\mathbf{A}_{n-1},\mathbf{E}_{n-1}).$$

为要计算这个行列式，我们交换它的最后两行使得行列式变号 (看第 188 页)．这样，最后一列的 $\mathbf{E}_{n-1}$ 就变成了 $\mathbf{E}_n$，因而由我们前面的结果就发现，$-z_{n-1}$ 等于从新矩阵中去掉最后一行最后一列所得到的行列式，或者等价地说，它等于原来方阵中 $a_{n-1,n}$ 的余子式．类似地，对于每一个 $i\ (=1,2,\cdots,n)$ 而言，$\pm z_i$ 等于元素 a_{i_n} 的余子式，其中正号对应着 $n-i$ 为偶数，负号对应着 $n-i$ 为奇数．

这样，公式 (75) 就成为一个 n 阶行列式按其最后一列元素及其对应的 $(n-1)$ 阶余子式相乘相加组成的展开式．例如对于 $n=4$，我们有

$$\begin{aligned}&\begin{vmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\\a_{41}&a_{42}&a_{43}&a_{44}\end{vmatrix}\\&=-a_{14}\begin{vmatrix}a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\\a_{41}&a_{42}&a_{43}\end{vmatrix}+a_{24}\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{31}&a_{32}&a_{33}\\a_{41}&a_{42}&a_{43}\end{vmatrix}\\&\quad-a_{34}\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{41}&a_{42}&a_{43}\end{vmatrix}+a_{44}\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}\end{aligned}\tag{77b}$$

适当作列与列之间的交换，我们能够得到一个行列式按其任意一个给定列的元素的余子式的展开式．正如我们在下节将要看到

的，这种类型的展开式在很多牵涉到空间维数的归纳法的证明中都起了作用.

c. 高维空间中的平行四边形的面积与平行多面体的体积

空间曲面能看成是由无限小的平行四边形构成的. 譬如曲面面积公式和曲面积分公式就要求有空间平行四边形面积的表达式的知识. 类似地，体积公式或流形上的体积积分公式必须建立在高维平行多面体体积的表达式上. 这种表达式能容易地借助于行列式以最大的一般性推导出来.

同向量有联系的基本量是两个向量

$$\mathbf{A}=(a_1,a_2,\cdots,a_n) \text{ 和 } \mathbf{B}=(b_1,b_2,\cdots,b_n)$$

的数量积，这在笛卡儿坐标系中由下式给出：

$$\mathbf{A}\cdot\mathbf{B}=a_1b_1+a_2b_2+\cdots+a_nb_n.$$

虽然 $\mathbf{A}$ 与 $\mathbf{B}$ 的各个分量 a_j 与 b_k 都依赖于所用的笛卡儿坐标系，但数量积却具有一种独立的**几何**意义：

$$\mathbf{A}\cdot\mathbf{B}=|\mathbf{A}||\mathbf{B}|\cos\gamma,$$

其中 $|\mathbf{A}|,|\mathbf{B}|$ 分别是向量 $\mathbf{A},\mathbf{B}$ 的长度，而 γ 是它们之间的夹角. 由此推出任何能表成数量积的量具有一种**不变的**几何意义，因而与所使用的特殊的笛卡儿坐标系无关.

能用数量积表示的最简单的量是两点 P_0, P_1 之间的距离，它是向量 $\mathbf{A}=\overrightarrow{P_0P_1}$ 的长度. 该距离的平方就是

$$|\mathbf{A}|^2=\mathbf{A}\cdot\mathbf{A}. \tag{78a}$$

对于 n 维空间中的两个向量 $\mathbf{A}$ 与 $\mathbf{B}$，如果给定它们以一个公共的起点 P_0，我们就能把它们联系到它们**所张成的平行四边形的面积**. 设 $\mathbf{A}=\overrightarrow{P_0P_1}$, $\mathbf{B}=\overrightarrow{P_0P_2}$. 则它们张成一个平行四边形 $P_0P_1QP_2$，以

P_1 和 P_2 作为顶点 P_0 的相邻顶点. 由初等几何知, 这平行四边形的面积 α 等于邻边长度的乘积乘以它们夹角的正弦

$$\alpha = |\mathbf{A}||\mathbf{B}|\sin\gamma = \sqrt{|\mathbf{A}|^2 \cdot |\mathbf{B}|^2 - |\mathbf{A}|^2|\mathbf{B}|^2\cos^2\gamma}$$
$$= \sqrt{|\mathbf{A}|^2|\mathbf{B}|^2 - (\mathbf{A}\cdot\mathbf{B})^2},$$

如同对于 $n = 3$ 的特殊情形我们在第 195 页求得的一样. 对于面积 α 的这个公式, 我们能用行列式的形式把 α 的平方写得更漂亮些

$$\alpha^2 = (\mathbf{A}\cdot\mathbf{A})(\mathbf{B}\cdot\mathbf{B}) - (\mathbf{A}\cdot\mathbf{B})(\mathbf{B}\cdot\mathbf{A}) = \begin{vmatrix} \mathbf{A}\cdot\mathbf{A} & \mathbf{A}\cdot\mathbf{B} \\ \mathbf{B}\cdot\mathbf{A} & \mathbf{B}\cdot\mathbf{B} \end{vmatrix}, \quad (78b)$$

上式右边出现的行列式叫作向量 $\mathbf{A},\mathbf{B}$ 的 **格拉姆行列式** 并且用 $\Gamma(\mathbf{A},\mathbf{B})$ 表示. 从推导清楚地有

$$\Gamma(\mathbf{A},\mathbf{B}) \geq 0$$

对于所有的向量 $\mathbf{A},\mathbf{B}$ 成立, 并且仅当 $\mathbf{A},\mathbf{B}$ 相关时等号成立 [1].

对于 n 维空间中 **三个向量** $\mathbf{A},\mathbf{B},\mathbf{C}$ **张成的平行六面体的体积**的平方, 我们能推导出类似的表达式. 我们将向量表成如下形式

$$\mathbf{A} = \overrightarrow{P_0P_1}, \quad \mathbf{B} = \overrightarrow{P_0P_2}, \quad \mathbf{C} = \overrightarrow{P_0P_3},$$

并且考虑以 P_1, P_2, P_3 为 P_0 相邻顶点的平行六面体. 它的体积 V 能够定义为它的一个面的面积 α 与对应高线长 h 的乘积. 选择 α 是向量 $\mathbf{A},\mathbf{B}$ 所张成的平行四边形的面积, h 表示点 P_3 到经过 P_0, P_1, P_2 的平面的距离. 这样就有

$$V^2 = h^2\alpha^2 = h^2\Gamma(A,B) = h^2\begin{vmatrix} \mathbf{A}\cdot\mathbf{A} & \mathbf{A}\cdot\mathbf{B} \\ \mathbf{B}\cdot\mathbf{A} & \mathbf{B}\cdot\mathbf{B} \end{vmatrix}.$$

我们说 h 表示 P_3 到平面 $P_0P_1P_2$ 的 "垂直" 距离, 也就是, 垂直于平面并且有终点 P 在平面上的向量 $\mathbf{D} = \overrightarrow{PP_3}$ 的长度. 对于平面

1) 那就是说, 如果有一个为零向量 ($\mathbf{A}$ 或 $\mathbf{B} = 0$) 或如果它们是平行的 ($\sin\gamma = 0$).

$P_0P_1P_2$ 上一点 P, 向量 $\overrightarrow{P_0P}$ 必定相关于 $\mathbf{A}=\overrightarrow{P_0P_1}$ 和 $\mathbf{B}=\overrightarrow{P_0P_2}$(看第 149 页),

$$\overrightarrow{P_0P}=\lambda\mathbf{A}+\mu\mathbf{B}.$$

因此，向量 $\mathbf{D}$ 有如下形式

$$\mathbf{D}=\overrightarrow{PP_3}=\overrightarrow{P_0P_3}-\overrightarrow{P_0P}=\mathbf{C}-\lambda\mathbf{A}-\mu\mathbf{B},$$

其中 λ 和 μ 为适当的常数．如果向量 $\mathbf{D}$ 垂直于向量 $\mathbf{A}$ 和 $\mathbf{B}$ 所张成的平面，则必有

$$\mathbf{A}\cdot\mathbf{D}=0,\quad \mathbf{B}\cdot\mathbf{D}=0. \tag{79a}$$

这就引出确定 λ 与 μ 的线性方程组

$$\mathbf{A}\cdot\mathbf{C}=\lambda\mathbf{A}\cdot\mathbf{A}+\mu\mathbf{A}\cdot\mathbf{B},\ \mathbf{B}\cdot\mathbf{C}=\lambda\mathbf{B}\cdot\mathbf{A}+\mu\mathbf{B}\cdot\mathbf{B}. \tag{79b}$$

这个方程组的行列式正好是格拉姆行列式 $\Gamma(\mathbf{A},\mathbf{B})$．假定 $\mathbf{A}$ 与 $\mathbf{B}$ 是无关的向量，我们就有 $\Gamma(\mathbf{A},\mathbf{B})\neq 0$．于是方程组 (79b) 存在唯一一组确定的解 λ,μ．因此，存在唯一的向量 $\mathbf{D}=\overrightarrow{PP_3}$ 垂直于平面 $P_0P_1P_2$ 且起点在该平面上．这个向量的长度等于距离 h, 所以由 (79a) 有

$$\begin{aligned}h^2=|\mathbf{D}|^2=\mathbf{D}\cdot\mathbf{D}&=(\mathbf{C}-\lambda\mathbf{A}-\mu\mathbf{B})\cdot\mathbf{D}\\&=\mathbf{C}\cdot\mathbf{D}-\lambda\mathbf{A}\cdot\mathbf{D}-\mu\mathbf{B}\cdot\mathbf{D}\\&=\mathbf{C}\cdot\mathbf{D}=\mathbf{C}\cdot\mathbf{C}-\lambda\mathbf{C}\cdot\mathbf{A}-\mu\mathbf{C}\cdot\mathbf{B}.\end{aligned}$$

由此得出

$$V^2=(\mathbf{C}\cdot\mathbf{C}-\lambda\mathbf{A}\cdot\mathbf{C}-\mu\mathbf{B}\cdot\mathbf{C})\Gamma(\mathbf{A},\mathbf{B}). \tag{79c}$$

这个关于 $\mathbf{A},\mathbf{B},\mathbf{C}$ 张成的平行六面体体积平方的表达式能更漂亮地写成由向量 $\mathbf{A},\mathbf{B},\mathbf{C}$ 作成的格拉姆行列式：

$$V^2=\begin{vmatrix}\mathbf{A}\cdot\mathbf{A} & \mathbf{A}\cdot\mathbf{B} & \mathbf{A}\cdot\mathbf{C}\\ \mathbf{B}\cdot\mathbf{A} & \mathbf{B}\cdot\mathbf{B} & \mathbf{B}\cdot\mathbf{C}\\ \mathbf{C}\cdot\mathbf{A} & \mathbf{C}\cdot\mathbf{B} & \mathbf{C}\cdot\mathbf{C}\end{vmatrix}=\Gamma(\mathbf{A},\mathbf{B},\mathbf{C}). \tag{79d}$$

为要证明关于 V^2 的表达式 (79d) 恒等于 (79c), 我们利用这个事实, 即如果在行列式 $\Gamma(\mathbf{A},\mathbf{B},\mathbf{C})$ 中把最后一列减去第一列的 λ 倍和第二列的 μ 倍, 则行列式的值不变:

$$\Gamma(\mathbf{A},\mathbf{B},\mathbf{C}) = \begin{vmatrix} \mathbf{A}\cdot\mathbf{A} & \mathbf{A}\cdot\mathbf{B} & \mathbf{A}\cdot\mathbf{C}-\lambda\mathbf{A}\cdot\mathbf{A}-\mu\mathbf{A}\cdot\mathbf{B} \\ \mathbf{B}\cdot\mathbf{A} & \mathbf{B}\cdot\mathbf{B} & \mathbf{B}\cdot\mathbf{C}-\lambda\mathbf{B}\cdot\mathbf{A}-\mu\mathbf{B}\cdot\mathbf{B} \\ \mathbf{C}\cdot\mathbf{A} & \mathbf{C}\cdot\mathbf{B} & \mathbf{C}\cdot\mathbf{C}-\lambda\mathbf{C}\cdot\mathbf{A}-\mu\mathbf{C}\cdot\mathbf{B} \end{vmatrix}.$$

由 (79b) 推出

$$\Gamma(\mathbf{A},\mathbf{B},\mathbf{C}) = \begin{vmatrix} \mathbf{A}\cdot\mathbf{A} & \mathbf{A}\cdot\mathbf{B} & 0 \\ \mathbf{B}\cdot\mathbf{A} & \mathbf{B}\cdot\mathbf{B} & 0 \\ \mathbf{C}\cdot\mathbf{A} & \mathbf{C}\cdot\mathbf{B} & \mathbf{C}\cdot\mathbf{C}-\lambda\mathbf{C}\cdot\mathbf{A}-\mu\mathbf{C}\cdot\mathbf{B} \end{vmatrix}.$$

按最后一列展开这个行列式, 立刻引回到 (79c).

公式 (79d) 表明, 向量 $\mathbf{A},\mathbf{B},\mathbf{C}$ 张成的平行六面体的体积 V 不依赖于计算中所选用的面和对应的高度, 因为我们交换 $\mathbf{A},\mathbf{B},\mathbf{C}$ 时 $\Gamma(\mathbf{A},\mathbf{B},\mathbf{C})$ 的值不变. 例如, $\Gamma(\mathbf{B},\mathbf{A},\mathbf{C})$ 能通过交换 $\Gamma(\mathbf{A},\mathbf{B},\mathbf{C})$ 的前两行和前两列得到.

公式 (79c) 能够写成

$$\Gamma(\mathbf{A},\mathbf{B},\mathbf{C}) = |\mathbf{D}|^2\Gamma(\mathbf{A},\mathbf{B}).$$

由此得到

$$\Gamma(\mathbf{A},\mathbf{B},\mathbf{C}) \geq 0$$

对于任何向量 $\mathbf{A},\mathbf{B},\mathbf{C}$ 成立. 这里等号仅仅在 $\Gamma(\mathbf{A},\mathbf{B}) = 0$ 或 $\mathbf{D} = \mathbf{O}$ 时成立. 但关系 $AJ = \Gamma(\mathbf{A},\mathbf{B}) = 0$ 隐含着 $\mathbf{A}$ 与 $\mathbf{B}$ 是相关的; 而如果 $\mathbf{D} = \mathbf{O}$, 则将有 $\mathbf{C} = \lambda\mathbf{A} + \mu\mathbf{B}$, 因而 $\mathbf{C}$ 将与 $\mathbf{A},\mathbf{B}$ 相关. 因此 **当且仅当向量** $\mathbf{A},\mathbf{B},\mathbf{C}$ **相关时, 格拉姆行列式等于零**.

对于 $n = 3$, 由三维空间中三个向量 $\mathbf{A},\mathbf{B},\mathbf{C}$ 所张成的平行六面体的体积 V 的公式 (74c) 立即推出公式 (79d). 这是第 173 页恒等式 (68f) 的推论, 根据它有

$$\det(\mathbf{A},\mathbf{B},\mathbf{C})\det(\mathbf{A},\mathbf{B},\mathbf{C}) = \Gamma(\mathbf{A},\mathbf{B},\mathbf{C}).$$

把 V^2 表示成格拉姆行列式，有益于表明 V 与所使用的特殊的笛卡儿坐标系无关，因而 V 具有一种几何意义.

我们还能讨论 n 维 $(n \geq 4)$ 空间中四个向量

$$\mathbf{A} = \overrightarrow{P_0P_1}, \quad \mathbf{B} = \overrightarrow{P_0P_2}, \quad \mathbf{C} = \overrightarrow{P_0P_3}, \quad \mathbf{D} = \overrightarrow{P_0P_4}$$

所张成的平行多面体的“体积” V. 定义 V 为向量 $\mathbf{A},\mathbf{B},\mathbf{C}$ 所张的三维平行六面体的体积与由点 P_4 到经过 P_0,P_1,P_2,P_3 的三维“平面”的距离相乘的乘积，经过与前面完全相同的步骤，我们就得到关于 V^2 的一个作为格拉姆行列式的表达式

$$\begin{aligned} V^2 &= \begin{vmatrix} \mathbf{A}\cdot\mathbf{A} & \mathbf{A}\cdot\mathbf{B} & \mathbf{A}\cdot\mathbf{C} & \mathbf{A}\cdot\mathbf{D} \\ \mathbf{B}\cdot\mathbf{A} & \mathbf{B}\cdot\mathbf{B} & \mathbf{B}\cdot\mathbf{C} & \mathbf{B}\cdot\mathbf{D} \\ \mathbf{C}\cdot\mathbf{A} & \mathbf{C}\cdot\mathbf{B} & \mathbf{C}\cdot\mathbf{C} & \mathbf{C}\cdot\mathbf{D} \\ \mathbf{D}\cdot\mathbf{A} & \mathbf{D}\cdot\mathbf{B} & \mathbf{D}\cdot\mathbf{C} & \mathbf{D}\cdot\mathbf{D} \end{vmatrix} \\ &= \Gamma(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}). \end{aligned} \tag{80a}$$

如果这里 $n = 4$, 这个格拉姆行列式就变成以 $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$ 为列向量的行列式的平方，从而我们得到

$$V = |\det(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D})|. \tag{80b}$$

更一般地，当指定一个公共的起点时，n 维空间中 m 个向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 张成一个 m 维的平行多面体，这个平行多面体的体积 V 的平方可用格拉姆行列式写成

$$\begin{aligned} V^2 &= \begin{vmatrix} \mathbf{A}_1\cdot\mathbf{A}_1 & \mathbf{A}_1\cdot\mathbf{A}_2 & \cdots & \mathbf{A}_1\cdot\mathbf{A}_m \\ \mathbf{A}_2\cdot\mathbf{A}_1 & \mathbf{A}_2\cdot\mathbf{A}_2 & \cdots & \mathbf{A}_2\cdot\mathbf{A}_m \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_m\cdot\mathbf{A}_1 & \mathbf{A}_m\cdot\mathbf{A}_2 & \cdots & \mathbf{A}_m\cdot\mathbf{A}_m \end{vmatrix} \\ &= \Gamma(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m). \end{aligned} \tag{81a}$$

对于 $m = n$ 我们得到关于 n 维空间中 n 个向量张成的平行多面体的体积 V 的公式

$$V = |\det(\mathbf{A}_1, \cdots, \mathbf{A}_n)|. \tag{81b}$$

可对 m 用归纳法证明

$$\Gamma(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m) \geq 0,$$

当且仅当 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 是相关时，上式等号成立[1].

d. n 维空间中平行多面体的定向

以后，在第五章，当我们需要一个相容的方法来确定多重积分的符号时，我们必须使用带符号的体积和 n 维空间中平行多面体的定向.

对于 n 维空间中 n 个向量所张的体积，我们有 (81b) 所给出的表达式

$$V = |\det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n)|.$$

我们称 $\det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n)$ 为在 $(x_1, \cdots, x_n)$ 坐标系中由 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 张成的平行多面体的体积. 如果 $\det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n)$ 是正的，就称这个平行多面体或向量组 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 关于坐标系有正的定向；如果行列式是负的，就说是有负的定向. 这样便有

$$\det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n) = \varepsilon V, \tag{81c}$$

其中 V 是由向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 张成的平行多面体的体积，而 $\varepsilon = +1$ 或 -1 则取决于平行多面体关于坐标系具有正的定向或负的定向.

尽管 $\det(\mathbf{A}_1, \cdots, \mathbf{A}_n)$ 的平方具有与笛卡儿坐标系无关的几何意义，但对于行列式的符号则并非如此. 例如，交换 x_1 轴与 x_2 轴的结果是交换行列式的前两行，因此 $\det(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n)$ 改变符号. 然而一种独立的几何意义所具有的则是：n 维空间中**两个**平行多面体总是或者具有**相同**的定向，或者具有**相反**的定向.

---

1) 在向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_m$ 相关的情形，它们以一点 P_0 为公共起点时所张成的平行多面体这时"退化"为一个 $m-1$ 维或者是低于 m 维的线性流型，因而这时 m 维体积等于零.

考虑 n 维空间中两个有序向量组 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 和 $\mathbf{B}_1, \mathbf{B}_2, \cdots, \mathbf{B}_n$，假定每一组都是由无关的向量组成的. 显然，当且仅当条件

$$\det(\mathbf{A}_1, \cdots, \mathbf{A}_n) \cdot \det(\mathbf{B}_1, \cdots, \mathbf{B}_n) > 0 \tag{82a}$$

满足时，这两个组向量才有相同的定向 —— 就是说，关于坐标系 $(x_1, x_2, \cdots, x_n)$ 二者都有正的定向或者都有负的定向. 利用等式 (68f)，我们能把条件 (82a) 写成如下形式

$$[\mathbf{A}_1, \cdots, \mathbf{A}_n; \mathbf{B}_1, \cdots, \mathbf{B}_n] > 0, \tag{82b}$$

其中左边的符号表示 $2n$ 个向量的函数，定义为

$$[\mathbf{A}_1, \cdots, \mathbf{A}_n; \mathbf{B}_1, \cdots, \mathbf{B}_n] = \begin{vmatrix} \mathbf{A}_1 \cdot \mathbf{B}_1 & \mathbf{A}_1 \cdot \mathbf{B}_2 & \cdots & \mathbf{A}_1 \cdot \mathbf{B}_n \\ \mathbf{A}_2 \cdot \mathbf{B}_1 & \mathbf{A}_2 \cdot \mathbf{B}_2 & \cdots & \mathbf{A}_2 \cdot \mathbf{B}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_n \cdot \mathbf{B}_1 & \mathbf{A}_n \cdot \mathbf{B}_2 & \cdots & \mathbf{A}_n \cdot \mathbf{B}_n \end{vmatrix}. \tag{82c}$$

注意当 $\mathbf{B}_1 = \mathbf{A}_1, \mathbf{B}_2 = \mathbf{A}_2, \cdots, \mathbf{B}_n = \mathbf{A}_n$ 时，符号 $[\mathbf{A}_1, \cdots, \mathbf{A}_n; \mathbf{B}_1, \cdots, \mathbf{B}_n]$ 简化为格拉姆行列式 $\Gamma(\mathbf{A}_1, \cdots, \mathbf{A}_n)$. 公式 (82b, c) 明显地表明：具有相同的定向是一种与所使用的特定的笛卡儿坐标系无关的几何性质. 我们可用符号

$$\Omega(\mathbf{A}_1, \cdots, \mathbf{A}_n) = \Omega(\mathbf{B}_1, \cdots, \mathbf{B}_n) \tag{82d}$$

表示这个性质，并且用

$$\Omega(\mathbf{A}_1, \cdots, \mathbf{A}_n) = -\Omega(\mathbf{B}_1, \cdots, \mathbf{B}_n) \tag{82e}$$

表示有相反的定向 [1]. 这样就可以更一般地说，对于 n 维空间中，

1) Ω 是个 n 重向量的一个定向排列，并不是一个"数". 公式 (82d, e) 表示定向的相等或不相等，公式 (82f) 仅仅把值 ± 1 与两个定向的比结合起来. 当然，完全可能用数值来描写 n 重向量两个不同的可能的定向，比如，对一个定向给出值 $\Omega = +1$，而对另一个定向给出值 $\Omega = -1$. 但是，这隐含着一个"标准定向"的任意选择. 例如我们称由坐标向量组给出的为 $+1$ —— 而关系 (82d, e, f) 却有着与 Ω 的任意指定值无关的意义. 类似的情况在数学中经常遇到. 例如，在欧几里得几何中，即使没有对距离赋以数值（如在欧几里得《几何原本》中）时，距离的相等甚到距离的比都有意义. 确实，我们能用实数描写距离，使得距离之比正是对应的实数之比. 这就要求任意选定一个"标准距离"(例如，一米)，以它为准可导出所有其他的距离，而这就在某种意义下引进了一种"非几何的"元素.

两组无关的向量有

$$\begin{aligned}&\Omega(\mathbf{B}_1,\cdots,\mathbf{B}_n)\\&=\operatorname{sgn}[\mathbf{A}_1,\cdots,\mathbf{A}_n;\mathbf{B}_1,\cdots,\mathbf{B}_n]\Omega(\mathbf{A}_1,\cdots,\mathbf{A}_n).\end{aligned}\tag{82f}$$

向量组 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n$ 关于 $(x_1,\cdots,x_n)$ 坐标系是正的定向或负的定向要按照

$$\Omega(\mathbf{A}_1,\cdots,\mathbf{A}_n)=\Omega(\mathbf{E}_1,\cdots,\mathbf{E}_n)\tag{83a}$$

或

$$\Omega(\mathbf{A}_1,\cdots,\mathbf{A}_n)=-\Omega(\mathbf{E}_1,\cdots,\mathbf{E}_n)\tag{83b}$$

二者来确定，其中 $\mathbf{E}_1,\cdots,\mathbf{E}_n$ 是坐标向量. 有时，我们将要用

$$\Omega(x_1,\cdots,x_n)$$

表示坐标系的定向 $\Omega(\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_n)$. 对于 n 维空间中两组 n 个向量的集合 $\mathbf{A}_1,\cdots,\mathbf{A}_n$ 和 $\mathbf{A}'_1,\cdots,\mathbf{A}'_n$, 由 (82c), (81b), 我们有

$$[\mathbf{A}_1,\cdots,\mathbf{A}_n;\mathbf{A}'_1,\cdots,\mathbf{A}'_n]=\varepsilon\varepsilon' VV',\tag{84a}$$

其中 V 和 V' 分别表示这两组向量所张成的平行多面体的体积; 因子 ε,ε' 依赖于它们以及坐标向量的定向:

$$\varepsilon=\operatorname{sgn}[\mathbf{A}_1,\cdots,\mathbf{A}_n;\mathbf{E}_1,\cdots,\mathbf{E}_n],\tag{84b}$$

$$\varepsilon'=\operatorname{sgn}[\mathbf{A}'_1,\cdots,\mathbf{A}'_n;\mathbf{E}_1,\cdots,\mathbf{E}_n].\tag{84c}$$

乘积

$$\varepsilon\varepsilon'=\operatorname{sgn}[\mathbf{A}_1,\cdots,\mathbf{A}_n;\mathbf{A}'_1,\cdots,\mathbf{A}'_n]$$

却与坐标系的选择无关, 并且当平行多面体有同样的定向时其值为 +1; 有相反的定向时其值为 −1.

利用数量积的定义, 对于 n 维空间中任意 $2m$ 个向量 $\mathbf{A}_1, \cdots, \mathbf{A}'_m$, 我们能作出表达式

$$[\mathbf{A}_1, \cdots, \mathbf{A}_m; \mathbf{A}'_1, \cdots, \mathbf{A}'_m] = \begin{vmatrix} \mathbf{A}_1 \cdot \mathbf{A}'_1 & \mathbf{A}_1 \cdot \mathbf{A}'_2 & \cdots & \mathbf{A}_1 \cdot \mathbf{A}'_m \\ \mathbf{A}_2 \cdot \mathbf{A}'_1 & \mathbf{A}_2 \cdot \mathbf{A}'_2 & \cdots & \mathbf{A}_2 \cdot \mathbf{A}'_m \\ \vdots & \vdots & & \vdots \\ \mathbf{A}_m \cdot \mathbf{A}'_1 & \mathbf{A}_m \cdot \mathbf{A}'_2 & \cdots & \mathbf{A}_m \cdot \mathbf{A}'_m \end{vmatrix} \tag{85a}$$

由定义很清楚, 这个表达式是 $2m$ 个向量的多线性型. 例如, 向量 $\mathbf{A}'_1$ 只出现在第一列, 并且该列元素是 $\mathbf{A}'_1$ 的线性型. 因为整个行列式是第一列元素的线性型, 可见它是 $\mathbf{A}'_1$ 的线性型. 由 (85a), 这个表达式对于固定的 $\mathbf{A}_1, \cdots, \mathbf{A}_m$, 它显然也是 $\mathbf{A}'_1, \cdots, \mathbf{A}'_m$ 的交替函数; 并且对于固定的 $\mathbf{A}'_1, \cdots, \mathbf{A}'_m$, 它是 $\mathbf{A}_1, \cdots, \mathbf{A}_m$ 的交替函数. 于是有 (看第 211 页的脚注)

$$[\mathbf{A}_1, \cdots, \mathbf{A}_m; \mathbf{A}'_1, \cdots, \mathbf{A}'_m] = 0. \tag{85b}$$

当这 m 个向量 $\mathbf{A}_1, \cdots, \mathbf{A}_m$ 或这 m 个向量 $\mathbf{A}'_1, \cdots, \mathbf{A}'_m$ 是相关的时候. 特别是, 当 $m > n$ 时 (85b) 永远成立.

于是, 我们假定 $m \leq n$, 并且假定这两组向量 $\mathbf{A}_1, \cdots, \mathbf{A}_m$ 和 $\mathbf{A}'_1, \cdots, \mathbf{A}'_m$ 都是无关的. 我们能够假定所有这些向量都给定了相同的起点, 譬如说就是 n 维空间的原点. 于是 $\mathbf{A}_1, \cdots, \mathbf{A}_m$ 张成一个经过原点的 m 维线性流型 π, 并且 $\mathbf{A}'_1, \cdots, \mathbf{A}'_m$ 张成另外一个这样的流形 π'. 在 π 中引进一组标准正交向量系 $\mathbf{E}_1, \cdots, \mathbf{E}_m$ 作为坐标向量, 并且在 π' 中引进另一组标准正交向量系 $\mathbf{E}'_1, \cdots, \mathbf{E}'_m$. [1)]对于固定的 $\mathbf{A}_1, \cdots, \mathbf{A}_m$, 函数 (85b) 是 $\mathbf{A}'_1, \cdots, \mathbf{A}'_m$ 的一个多线性交替型, 因而 (看 162 页) 有

$$\begin{aligned} &[\mathbf{A}_1, \cdots, \mathbf{A}_m; \mathbf{A}'_1, \cdots, \mathbf{A}'_m] \\ =&[\mathbf{A}_1, \cdots, \mathbf{A}_m; \mathbf{E}'_1, \cdots, \mathbf{E}'_m] \det(\mathbf{A}'_1, \cdots, \mathbf{A}'_m) \end{aligned}$$

1) π 和 π' 这两组坐标向量, 不一定以任何方式互相联系, 也不与包含 π 和 π' 的整个 n 维空间的坐标系有什么关系.

其中 $\det(\mathbf{A}'_1,\cdots,\mathbf{A}'_m)$ 是向量 $\mathbf{A}'_1,\cdots,\mathbf{A}'_m$ 关于坐标向量 $\mathbf{E}'_1,\cdots,\mathbf{E}'_m$ 的分量所构成的方阵的行列式. 系数 $[\mathbf{A}_1,\cdots,\mathbf{A}_m;\mathbf{E}'_1,\cdots,\mathbf{E}'_m]$ 本身显然是 $\mathbf{A}_1,\cdots,\mathbf{A}_m$ 的一个多线性交替型, 因而等于

$$[\mathbf{E}_1,\cdots,\mathbf{E}_m;\mathbf{E}'_1,\cdots,\mathbf{E}'_m]\det(\mathbf{A}_1,\cdots,\mathbf{A}_m),$$

其中最后一个行列式是从向量 $\mathbf{A}_1,\cdots,\mathbf{A}_m$ 关于坐标向量 $\mathbf{E}_1,\cdots,\mathbf{E}_m$ 的分量的方阵的行列式构成的.

利用公式 (81c), 我们得到恒等式

$$[\mathbf{A}_1,\cdots,\mathbf{A}_m;\mathbf{A}'_1,\cdots,\mathbf{A}'_m]=\mu\varepsilon\varepsilon' VV', \tag{85c}$$

其中 V 和 V' 分别是由向量 $\mathbf{A}_1,\cdots,\mathbf{A}_m$ 和 $\mathbf{A}'_1,\cdots,\mathbf{A}'_m$ 张成的平行多面体的体积; 因子 ε,ε' 则把这两个平行多面体的定向联系到 π 和 π' 的坐标系的定向, 就是

$$\varepsilon=\operatorname{sgn}[\mathbf{A}_1,\cdots,\mathbf{A}_m;\mathbf{E}_1,\cdots,\mathbf{E}_m],$$
$$\varepsilon'=\operatorname{sgn}[\mathbf{A}'_1,\cdots,\mathbf{A}'_m;\mathbf{E}'_1,\cdots,\mathbf{E}'_m].$$

最后, 系数

$$\mu=[\mathbf{E}_1,\cdots,\mathbf{E}_m;\mathbf{E}'_1,\cdots,\mathbf{E}'_m]$$

仅仅依赖于空间 π 和 π' 以及在这些空间中所选取的坐标系. 如 $\pi=\pi'$, 我们能选取

$$\mathbf{E}'_1=\mathbf{E}_1,\cdots,\mathbf{E}'_m=\mathbf{E}_m,$$

这时 $\mu=1$, 与 (84a) 一样.

对于 $\mu\neq 0$, 我们能用 (85c) 来联系 n 维空间中两个不同的 m 维线性流型 π 和 π' 的定向 [1]. 我们总是能假定 $\mu>0$, 因为如果需要, 就可把一个坐标向量换成与它相反的向量. 于是由 (85c) 有

$$\operatorname{sgn}[\mathbf{A}_1,\cdots,\mathbf{A}_m;\mathbf{A}'_1,\cdots,\mathbf{A}'_m]=\varepsilon\varepsilon'.$$

---

1) 容易验证, $\mu=0$ 仅仅当 π 与 π' 是**互相垂直**时成立, 这也就是说, π' 包含一个向量垂直于 π 中所有的向量. 更一般地, 系数 μ 能被解释为这两个流型之间的夹角的余弦 (看第 219 页问题 13).

因此，对于 π 中任意向量 $\mathbf{A}_1, \cdots, \mathbf{A}_m$ 和 π' 中任意向量 $\mathbf{A}'_1, \cdots, \mathbf{A}'_m$ 而言，条件

$$[\mathbf{A}_1, \cdots, \mathbf{A}_m; \mathbf{A}'_1, \cdots, \mathbf{A}'_m] > 0$$

意味着这两组向量关于 π 与 π' 这两个空间中的坐标系都有正的定向或都有负的定向.

e. 平面与超平面的定向

在一个 m 维线性流型 π 中，一个笛卡儿坐标系的选择决定着一个确定的方向

$$\Omega(\mathbf{E}_1, \cdots, \mathbf{E}_m),$$

其中 $\mathbf{E}_1, \cdots, \mathbf{E}_m$ 为坐标向量. 这种选择规定了 π 中哪些向量组 $\mathbf{A}_1, \cdots, \mathbf{A}_m$ 是被称为是有正定向的，即与 $\mathbf{E}_1, \cdots, \mathbf{E}_m$ 有相同定向的那些向量组. 我们用 π^* 表示线性空间 π 与其中选定的一个定向的组合，并且称 π^* 是一个**定向线性**流型. 我们把选定的定向记作 $\Omega(\pi^*)$，并且把 π 中 m 个无关的向量 $\mathbf{A}_1, \cdots, \mathbf{A}_m$ 称作是正向的，如果

$$\Omega(\mathbf{A}_1, \cdots, \mathbf{A}_m) = \Omega(\pi^*).$$

相对于 x^* 中一个笛卡儿坐标系，如果坐标向量与 π^* 有同样的定向，我们就称 π^* 是**正向的**.

一个定向的二维平面 π^*，可以被看作一个区别了**正旋转方向**的平面. 如果一对向量 $\mathbf{A}, \mathbf{B}$ 关于 π^* 有"正"的定向，则 π^* 的正的旋转方向便是将 $\mathbf{A}$ 的方向旋转一个小于 $180°$ 的角达到 $\mathbf{B}$ 的方向的那个定向 [1].

---

1) 注意 π^* 的定向只能由 π 中指出一个特定的正向向量对 $\mathbf{B}, \mathbf{C}$，或者由 π 中一个区别了旋转方向的旋转物 (如钟) 来描画. 没有抽象的方法来确定一个给定的旋转是**顺时针**还是**逆时针**，就像没有抽象的方法说出哪是左侧哪是右侧一样. 这些问题只能参考到某些标准对象来确定.

如果有向二维平面 π^* 位于有向三维平面 σ^* 之中，我们就能区分出 π^* 的 正 侧与 负 侧. 设 P_0 是 π^* 的任意一点，我们在 π^* 内取两个无关的向量 $\mathbf{B}=\overrightarrow{P_0P_1}$, $\mathbf{C}=\overrightarrow{P_0P_2}$ 使得

$$\Omega(\mathbf{B},\mathbf{C})=\Omega(\pi^*). \tag{86a}$$

与向量 $\mathbf{B},\mathbf{C}$ 无关的第三向量 $\mathbf{A}=\overrightarrow{P_0P_3}$ 称为指向 π^* 的 正侧，如果有

$$\Omega(\mathbf{A},\mathbf{B},\mathbf{C})=\Omega(\sigma^*). \tag{86b}$$

如果 σ^* 相对于一个笛卡儿坐标系有正的定向，则在这个坐标系中我们能用

$$\det(\mathbf{A},\mathbf{B},\mathbf{C})>0 \tag{86c}$$

代替 (86b). 如果 σ^* 相对于常用的右手坐标系有正的定向，则从有向平面 π^* 正的一侧看去，π^* 的正的旋转方向表现为逆时针方向.

同样的术语适用于 n 维有向空间 σ^* 中的有向超平面 π^*. 给定了 π^* 中 $n-1$ 个向量 $\mathbf{A}_2,\cdots,\mathbf{A}_n$, 使得

$$\Omega(\mathbf{A}_2,\cdots,\mathbf{A}_n)=\Omega(\pi^*), \tag{87}$$

我们就称一个向量 $\mathbf{A}_1$ 是指向 π^* 的正侧，如果

$$\Omega(\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n)=\Omega(\sigma^*).$$

f. 线性变换下平行多面体体积的改变

一个 n 行 n 列的方阵 $\mathbf{a}=(a_{ji})$ 确定了把 n 维空间的向量 $\mathbf{X}$ 变到该空间内的向量 $\mathbf{Y}$ 的一个线性变换或线性映射 $\mathbf{Y}=\mathbf{aX}$. 这里，我们假定 $\mathbf{X}$ 与 $\mathbf{Y}$ 参考于相同的坐标向量 $\mathbf{E}_1,\cdots,\mathbf{E}_n$. 对于 $\mathbf{X}=(x_1,\cdots,x_n)$, $\mathbf{Y}=(y_1,\cdots,y_n)$, 这变换可用分量写成如下形式

$$y_j=\sum_{r=1}^{n}a_{jr}x_r \quad (j=1,\cdots,n).$$

一组 n 个向量 $\mathbf{B}_1=(b_{11},\cdots,b_{n1}),\cdots,\mathbf{B}_n=(b_{1n},\cdots,b_{nn})$ 变换到另一组 n 个向量 $\mathbf{C}_1=(c_{11},\cdots,c_{n1}),\cdots,\mathbf{C}_n=(c_{1n},\cdots,c_{nn})$, 其中

$$c_{jk}=\sum_{r=1}^{n}a_{jr}b_{rk}\quad(j=1,2,\cdots,n).$$

根据矩阵乘法的行列式规则，我们有

$$\det(\mathbf{C}_1,\cdots,\mathbf{C}_n)=\det(\mathbf{a})\cdot\det(\mathbf{B}_1,\cdots,\mathbf{B}_n). \tag{88a}$$

这个公式包含了这样两个公式:

$$|\det(\mathbf{C}_1,\cdots,\mathbf{C}_n)|=|\det(\mathbf{a})||\det(\mathbf{B}_1,\cdots,\mathbf{B}_n)|, \tag{88b}$$

$$\operatorname{sgn}\det(\mathbf{C}_1,\cdots,\mathbf{C}_n)=[\operatorname{sgn}\det(\mathbf{a})][\operatorname{sgn}\det(\mathbf{B}_1,\cdots,\mathbf{B}_n)]. \tag{88c}$$

这两个规则可以直接用几何语言陈述如下:

对应于一个方阵 $\mathbf{a}$ 的, 把 n 维空间变换到它自身中的线性变换, 将每个由 n 个向量张成的平行多面体的体积乘上了相同的常数因子 $|\det(\mathbf{a})|$. 如果 $\det(\mathbf{a})>0$, 它就保持了所有平行多面体的定向, 如果 $\det(\mathbf{a})<0$, 它就改变了所有平行多面体的定向. [1)]

对于刚体运动，方阵 $\mathbf{a}$ 是正交的，因而 (看第 189 页) 它的行列式等于 $+1$ 或 -1. 因此，刚体运动保持平行多面体的体积; 对于 $\det(\mathbf{a})=+1$ 还保持了定向; 其他都改变了定向.

练　习　2.4

1. 利用向量积术语讨论练习 2.2 的第 5 题.

2. 在一个等速旋转中, 设旋转轴经过原点, 它的方向角是 α,β,γ, 旋转的角速度是 w, 求点 P 的速度.

---

1) 强调这个定理的假设是重要的. 仅仅是 n 维平行多面体的体积被乘上了这一共同的因子, 低维的多面体被乘上的因子随着它们的位置而变化. 还有, 如果要关于定向的叙述成立, 我们就不得不假定像与原像是参考于同一个坐标系的.

3. 证明经过三点 $(x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3)$ 的平面具有方程

$$\begin{vmatrix} x_1 - x & y_1 - y & z_1 - z \\ x_2 - x & y_2 - y & z_2 - z \\ x_3 - x & y_3 - y & z_3 - z \end{vmatrix} = 0.$$

4. 求空间中两直线 l 到 l' 之间的最短距离, l 由方程组 $x = at + b, y = ct + d, z = et + f$ 给出, l' 由方程组 $x = a't + b', y = c't + d', z = e't + f'$ 给出.

5. 试证以 $P_1(x_1, y_1), P_2(x_2, y_2), \cdots, P_n(x_n, y_n)$ 为相继顶点的凸多边形的面积是下列和数的绝对值的一半:

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} + \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} + \cdots + \begin{vmatrix} x_{n-1} & x_n \\ y_{n-1} & y_n \end{vmatrix} + \begin{vmatrix} x_n & x_1 \\ y_n & y_1 \end{vmatrix}.$$

6. 证明以 $(x_1, y_1), (x_2, y_2)$ 和 (x_3, y_3) 为顶点的三角形的 (有向) 面积是

$$\frac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}.$$

7. 试证明: 如果上题中三角形的顶点的坐标都是有理数, 则这三角形不能是等边的.

8. (a) 证明不等式

$$D = \begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix}$$
$$\leq \sqrt{(a^2 + b^2 + c^2)(a'^2 + b'^2 + c'^2)(a''^2 + b''^2 + c''^2)};$$

(b) 何时等号成立.

9. 证明向量恒等式

(a) $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$,

(b) $(\mathbf{X} \times \mathbf{Y}) \cdot (\mathbf{X}' \times \mathbf{Y}') = (\mathbf{X} \cdot \mathbf{X}')(\mathbf{Y} \cdot \mathbf{Y}') - (\mathbf{X} \cdot \mathbf{Y}')(\mathbf{Y} \cdot \mathbf{X}')$,

(c) $[\mathbf{X} \times (\mathbf{Y} \times \mathbf{Z})] \cdot \{[\mathbf{Y} \times (\mathbf{Z} \times \mathbf{X})] \times [\mathbf{Z} \times (\mathbf{X} \times \mathbf{Y})]\} = 0$.

10. 给出绕轴 $x : y : z = 1 : 0 : -1$ 旋转一个角度 ϕ 的变换公式, 使得从 $(-1, 0, 1)$ 看去平面 $x = z$ 的旋转是正的.

11. 如果 $\mathbf{A},\mathbf{B},\mathbf{C}$ 是无关的, 利用从练习 (9a) 得到的关于 $\mathbf{X}=(\mathbf{A}\times\mathbf{B})\times(\mathbf{C}\times\mathbf{D})$ 的两种表达式, 将 $\mathbf{D}$ 表成 $\mathbf{A},\mathbf{B},\mathbf{C}$ 的线性组合.

12. 设 Ox,Oy,Oz 和 Ox',Oy',Oz' 是两个右手坐标系. 假设 Oz 与 Oz' 不相重合; 设角 zOz' 是 θ $(0<\theta<\pi)$. 试引射线 Ox_1 使它与 Oz,Oz' 都成直角并且使 Ox_1,Oz,Oz' 系与 Ox,Oy,Oz 系有相同的定向. 这个 Ox_1 就是平面 Oxy 与平面 $Ox'y'$ 的交线. 设角 xOx_1 是 ϕ, 角 x_1Ox' 是 ψ, 并设它们在各自所属的平面 Oxy 与 $O'x'y'$ 中有通常的正的定向. 求坐标变换的矩阵.

13. 设 π 和 π' 是同一个 n 维空间中的两个 m 维线性子空间, 分别具有标准正交基 $\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_m$ 和 $\mathbf{E}'_1,\mathbf{E}'_2,\cdots,\mathbf{E}'_m$. 试证 $\mu=[\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_n;\mathbf{E}'_1,\mathbf{E}'_2,\cdots,\mathbf{E}'_m]=0$, 当且仅当 π 与 π' 是正交的, 即其中一个空间含有一个向量垂直于另一空间的所有向量.

2.5 分析中的向量概念

a. 向量场

当我们研究依赖一个或多个连续变动的参变量的**向量流形**时, 数学分析就开始起作用了.

例如, 如果我们考虑占据着空间的一部分并且处于运动状态的一个物体, 则在一个给定的时刻, 物体的每个质点都将有一个确定的由一个向量 $\mathbf{U}=(u_1,u_2,u_3)$ 表示的速度. 我们说这些向量在所讨论的区域中形成一个**向量场**. 于是, 这个场向量的三个分量就作为这质点在给定的时刻的位置的三个坐标的三个函数

$$u_1(x_1,x_2,x_3),\quad u_2(x_1,x_2,x_3),\quad u_3(x_1,x_2,x_3)$$

出现. 我们常常用一个以 (x_1,x_2,x_3) 为起点向量代表 $\mathbf{U}$.

作用在空间中不同点的力同样形成一个向量场. 作为一个**力场**的例子, 我们考虑一个重质点按照牛顿的引力定律对一个单位

质量的万有引力. 按照这个定律, 场向量 $\mathbf{F}=(f_1,f_2,f_3)$ 在每一点 (x_1,x_2,x_3) 处都指向吸引的质点, 并且它的大小反比于从质点到它的距离的平方.

场向量, 如 $\mathbf{U}$ 或 $\mathbf{F}$, 有独立于坐标系的物理意义. 在一个给定的笛卡儿 (x_1,x_2,x_3) 坐标系中, 向量 $\mathbf{U}$ 有依赖于坐标系的分量 u_1,u_2,u_3. 在另一个笛卡儿坐标系中, 原来具有坐标 x_1,x_2,x_3 的点有了坐标 y_1,y_2,y_3, 其中 y_j 与 x_k 由形如

$$\begin{cases} y_1=a_{11}x_1+a_{12}x_2+a_{13}x_3+b_1\\ y_2=a_{21}x_1+a_{22}x_2+a_{23}x_3+b_2\\ y_3=a_{31}x_1+a_{32}x_2+a_{33}x_3+b_3\end{cases} \tag{89a}$$

或

$$y_j=\sum_{k=1}^{3}a_{jk}x_k+b_j \quad (j=1,2,3) \tag{89b}$$

的方程组联系着. 于是向量 $\mathbf{U}$ 在新坐标系中的分量 v_1,v_2,v_3 便由齐次关系

$$v_j=\sum_{k=1}^{3}a_{jk}u_k \quad (j=1,2,3) \tag{89c}$$

给出.

方阵 $\mathbf{a}=(a_{jk})$ 是正交的, 所以 (见第 170 页) 它的逆等于它的转置. 因此, 方程组 (89b), (89c) 关于 x_k 与 u_k 的解取如下形式

$$x_k=\sum_{j=1}^{3}a_{jk}(y_j-b_j) \quad (k=1,2,3), \tag{89d}$$

$$u_k=\sum_{j=1}^{3}a_{jk}v_j \quad (k=1,2,3). \tag{89e}$$

变量 x_1,x_2,x_3 的任何三个函数 u_1,u_2,u_3 都确定一个向量场 $\mathbf{U}$. 在 (x_1,x_2,x_3) 坐标系中具有分量 u_1,u_2,u_3. 如果这个场要有一种独立于坐标系的几何意义, 则 $\mathbf{U}$ 在笛卡儿 (y_1,y_2,y_3) 坐标系中

的分量 v_i 就必定是由 (89c) 给出，只要 y_i 与 x_i 是由 (89a) 联系着.

b. 数量场的梯度

一个数量场是空间中点 P 的一个函数 $s=s(P)$. 在任何一个笛卡儿坐标系中，若点 P 由它的坐标 x_1,x_2,x_3 描写，则数量场 s 就变成一个函数 $s=f(x_1,x_2,x_3)$. 我们可以把这三个偏导数

$$u_1=\frac{\partial s}{\partial x_1}=f_{x_1}(x_1,x_2,x_3)$$
$$u_2=\frac{\partial s}{\partial x_2}=f_{x_2}(x_1,x_2,x_3),$$
$$u_3=\frac{\partial s}{\partial x_3}=f_{x_3}(x_1,x_2,x_3)$$

看作是一个向量 $\mathbf{U}=(u_1,u_2,u_3)$ 在 (x_1,x_2,x_3) 坐标系中的分量.

在任何一个与原来的坐标系有关系 (89a) 或 (89d) 的新的笛卡儿 (y_1,y_2,y_3) 坐标系中，数量场 s 就由函数

$$\begin{aligned}s&=g(y_1,y_2,y_3)\\&=f\left(\sum_{k=1}^{3}a_{k1}(y_k-b_k),\sum_{k=1}^{3}a_{k2}(y_k-b_k),\sum_{k=1}^{3}a_{k3}(y_k-b_k)\right)\end{aligned}$$

表示. 根据 *微分法的链式法则*(第 57 页), 我们有

$$v_j=\frac{\partial s}{\partial y_j}=g_{y_j}(y_1,y_2,y_3)=\sum_{k=1}^{3}\frac{\partial s}{\partial x_k}\cdot\frac{\partial x_k}{\partial y_j}=\sum_{k=1}^{3}u_k a_{jk}.$$

利用关系 (89c), 我们看到向量 $\mathbf{U}$ 在 (y_1,y_2,y_3) 坐标系中具有分量

$$v_j=\frac{\partial s}{\partial y_j}.$$

这样，数量 s 的偏导数就在任何一个笛卡儿坐标系中形成一个不依赖于坐标系的选择的一个向量 $\mathbf{U}$ 的分量. 我们称 $\mathbf{U}$ 为 *数量场 s 的梯度* 并且记作

$$\mathbf{U}=\operatorname{grad}s.$$

由第 50 页的公式 (14b), s 在具有方向余弦 $\cos\alpha_1, \cos\alpha_2, \cos\alpha_3$ 的方向上的方向导数在 (x_1, x_2, x_3) 坐标系中由下式给出:

$$D_{(\alpha)}s = \frac{\partial s}{\partial x_1}\cos\alpha_1 + \frac{\partial s}{\partial x_2}\cos\alpha_2 + \frac{\partial s}{\partial x_3}\cos\alpha_3. \tag{90a}$$

在带有方向角 $\alpha_1, \alpha_2, \alpha_3$ 的方向上引进单位向量 $\mathbf{R} = (\cos\alpha_1, \cos\alpha_2, \cos\alpha_3)$, 我们就能把 s 在该方向上的导数用向量记号写成

$$D_{(\alpha)}s = \mathbf{R} \cdot \operatorname{grad} s. \tag{90b}$$

由柯西 - 施瓦茨不等式 (见第 143 页), 对于 $|\mathbf{R}| = 1$, 我们得到

$$|D_{(\alpha)}s| \leq |\mathbf{R}||\operatorname{grad} s| = |\operatorname{grad} s|.$$

因此, **s 在任何一个方向上的导数都决不超过 s 的梯度的长度**. 取 $\mathbf{R}$ 为 $\operatorname{grad} s$ 方向上的单位向量, 我们就求得这个方向导数的值

$$D_{(\alpha)}s = \frac{1}{|\operatorname{grad} s|}(\operatorname{grad} s) \cdot (\operatorname{grad} s) = |\operatorname{grad} s|.$$

因此, **s 的梯度向量的长度就等于 s 在一切方向上的最大变化率. 梯度的方向是数量场增加最快的一个方向, 而在它的反方向上 s 减少得最快.**

我们将在第三章中回到梯度的几何解释. 然而, 我们立刻能给出梯度**方向**的一个直观概念. 我们首先限于考虑两维向量, 我们要考虑的是一个数量场 $s = f(x_1, x_2)$ 的梯度. 我们将假定 s 是由它在 (x_1, x_2) 平面内的等高线 (或等值线)

$$s = f(x_1, x_2) = 常数 = c$$

表出的. 则 s 在一点 P 处沿经过 P 点的等高线方向上的方向导数显然是 0, 因为如果 Q 点是该等高线上另外一点, 则

$$s(Q) - s(P) = 0$$

成立；用 Q 与 P 之间的距离 ρ 去除，再让 ρ 趋向于 0 求极限 (见第 48 页) 就得到 s 在 P 点沿等高线切线方向上的方向导数是 0. 因此，如果 $\mathbf{R}$ 是等高线的切线方向上的单位向量，就由 (90b) 有

$$\mathbf{R}\cdot\operatorname{grad}s=0;$$

所以 s 在每一点的梯度向量垂直于经过该点处的等高线. 在三维情形，对于梯度有完全类似的陈述成立. 如果我们把数量场 s 用它的等高面

$$s=f(x_1,x_2,x_3)=\text{常数}=c$$

表示，则梯度在等高面的每一个切线方向上的分量为零，因而它垂直于等高面.

在应用中，我们经常遇到代表一个数量函数的梯度的向量场. 集中在一点 $Q=(\xi_1,\xi_2,\xi_3)$ 的一个质量为 M 的质点产生的引力场可作为一个例子. 设 P 点处质量为 m 的质点受质量为 M 的引力为 $\mathbf{F}=(f_1,f_2,f_3)$. 用 $\mathbf{R}$ 表示向量

$$\mathbf{R}=\overrightarrow{QP}=(x_1-\xi_1,x_2-\xi_2,x_3-\xi_3).$$

由牛顿万有引力定律， $\mathbf{F}$ 的方向是 $-\mathbf{R}$, $\mathbf{F}$ 的大小是 $C/|\mathbf{R}|^2$, 其中 $C=\gamma mM$ (其中 γ 表示万有引力常数). 因此

$$\mathbf{F}=-\frac{C}{|\mathbf{R}|^3}\mathbf{R}$$

或

$$f_j=C\frac{\xi_j-x_j}{\sqrt{(\xi_1-x_1)^2+(\xi_2-x_2)^2+(\xi_3-x_3)^2}}\ (j=1\ 2,3).$$

由微分法可立即得到

$$f_j=\frac{\partial}{\partial x_j}\frac{C}{\sqrt{(\xi_1-x_1)^2+(\xi_2-x_2)^2+(\xi_3-x_3)^2}}\ (j=1,2,3).$$

因此

$$\mathbf{F} = \operatorname{grad} \frac{C}{\gamma}, \tag{91}$$

其中

$$\gamma = \sqrt{(\xi_1 - x_1)^2 + (\xi_2 - x_2)^2 + (\xi_3 - x_3)^2} = |\mathbf{R}|,$$

它是点 P 与 Q 之间的距离.

如果一个力场是一个数量函数的梯度，这个数量函数常常被称作这个场的**势函数**. 在功与能的研究中，我们将要从更一般的观点考虑这个概念.

c. 向量场的散度和旋度

通过微分法，我们已经对每一个数量场指定了一个向量场，即梯度. 类似地，通过微分法，我们能对每一个向量场 $\mathbf{U}$ 指定一个确定的数量场，称为向量场 $\mathbf{U}$ 的**散度**. 对于一个特定的笛卡儿 (x_1, x_2, x_3) 坐标系，在其中 $\mathbf{U} = (u_1, u_2, u_3)$，我们定义向量 $\mathbf{U}$ 的散度为这样一个函数

$$\operatorname{div} \mathbf{U} = \frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} + \frac{\partial u}{\partial x_3}, \tag{92}$$

也就是三个分量分别关于三个坐标的偏导数之和. 我们能证明，用这种方式定义的数量场 $\operatorname{div} \mathbf{U}$ 不依赖于特殊的笛卡儿坐标系的选择[1]. 设一点 (x_1, x_2, x_3) 在另一笛卡儿坐标系中的坐标是 y_1, y_2, y_3，两种坐标由方程组 (89b) 联系着；从而 $\mathbf{U}$ 在新坐标系中的分量由关系式 (89c) 给出. 由微分法的链式法则，我们有

$$\operatorname{div} \mathbf{U} = \sum_{k=1}^{3} \frac{\partial u_k}{\partial x_k} = \sum_{k,j=1}^{3} \frac{\partial u_k}{\partial y_j} \frac{\partial y_j}{\partial x_k}$$

---

1) 由向量 $\mathbf{U}$ 的分量的一阶微商所构成的其他表达式则不一定是这种情形，例如

$$\frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} - \frac{\partial u_3}{\partial x_3} \text{ 或 } \frac{\partial u_1}{\partial x_2} \cdot \frac{\partial u_2}{\partial x_3} \cdot \frac{\partial u_3}{\partial x_1}.$$

$$= \sum_{j,k=1}^{3} a_{jk} \frac{\partial u_k}{\partial y_j} = \sum_{j=1}^{3} \frac{\partial}{\partial y_j} \sum_{k=1}^{3} a_{jk} u_k$$
$$= \sum_{j=1}^{3} \frac{\partial v_j}{\partial y_j},$$

这表明在任何其他的坐标系中，我们得到相同的数量 $\operatorname{div} \mathbf{U}$.

这里我们满足于散度的形式定义, 它的物理意义将在以后讨论(第五章第 9 节).

对于一个向量场 $\mathbf{U}$ 的所谓*旋度*, 我们将采用同样的步骤. 旋度本身是一个向量

$$\mathbf{B} = \operatorname{curl} \mathbf{U}.$$

如果有一个 (x_1, x_2, x_3) 坐标系中，向量 $\mathbf{U}$ 具有分量 u_1, u_2, u_3, 我们就通过下式定义 curl $\mathbf{U}$ 的三个分量 b_1, b_2, b_3:

$$b_1 = \frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3}, \; b_2 = \frac{\partial u_1}{\partial x_3} - \frac{\partial u_3}{\partial x_1}, \; b_3 = \frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}. \tag{93}$$

同其他情形一样，我们能够验证关于向量 $\mathbf{U}$ 的旋度的定义实际上给出一个不依赖于特殊坐标系的选择的向量, 只要所考虑的这些笛卡儿坐标系有相同的定向. 但是这里我们省去了这些计算，因为在第五章中我们将给出旋度的物理解释, 它会清楚地显示出旋度的向量特征.

如果我们使用一个带有分量

$$\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}$$

的符号向量, 那么梯度、散度、旋度这三个概念就能互相联系起来. 这个*向量微分算子*常常用倒三角形的符号 ∇ 表示，读作 “del”. 数量场 s 的梯度是符号向量 ∇ 与数量 s 的乘积；就是说，它是这个向量 [1)]

$$\operatorname{grad} s = \nabla s = \left(\frac{\partial}{\partial x_1} s, \frac{\partial}{\partial x_2} s, \frac{\partial}{\partial x_3} s\right). \tag{94a}$$

1) 在乘积 ∇s 中我们不得不把向量写在数量的前面，这与我们通常的习惯相反，这是因为符号向量 ∇ 的分量不能与通常的数量交换.

向量场 $\mathbf{U}=(u_1,u_2,u_3)$ 的散度是数量积

$$\operatorname{div}\mathbf{U}=\nabla\cdot\mathbf{U}=\frac{\partial}{\partial x_1}u_1+\frac{\partial}{\partial x_2}u_2+\frac{\partial}{\partial x_3}u_3. \tag{94b}$$

最后，向量场 $\mathbf{U}$ 的旋度是向量积

$$\begin{aligned}\operatorname{curl}\mathbf{U}&=\nabla\times\mathbf{U} \qquad (94c)\\&=\left(\frac{\partial}{\partial x_2}u_3-\frac{\partial}{\partial x_3}u_2,\frac{\partial}{\partial x_3}u_1-\frac{\partial}{\partial x_1}u_3,\frac{\partial}{\partial x_1}u_2-\frac{\partial}{\partial x_2}u_1\right)\end{aligned}$$

[见第 196 页 (71b)]. 由微分法的链式法则，可以推出这个事实，即向量 ∇ 与用来定义它的分量的笛卡儿坐标系无关；在坐标变换 (89d) 之下，由链式法则我们有

$$\frac{\partial}{\partial y_j}=\sum_{k=1}^{3}\frac{\partial x_k}{\partial y_j}\frac{\partial}{\partial x_k}=\sum_{k=1}^{3}a_{jk}\frac{\partial}{\partial x_k},$$

这表明 ∇ 的分量按照向量的规则 (89c) 来变换. 由此明显看出 ∇s, $\nabla\cdot\mathbf{U}$, $\nabla\times\mathbf{U}$ 也不依赖于坐标系 [1].

最后，我们叙述一些经常出现的关系. **梯度的散度等于零**，写成符号便是

$$\operatorname{curl}\operatorname{grad}s=\nabla\times(\nabla s)=0. \tag{95a}$$

旋度的散度等于零，写成符号便是

$$\operatorname{div}\operatorname{curl}\mathbf{U}=\nabla\cdot(\nabla\times\mathbf{U})=0. \tag{95b}$$

我们容易看出，从梯度、散度、旋度的定义，应用微分法的可交换性便可推出这些关系. 如果对于符号向量 ∇ 应用向量的通常的规则，也可以形式地推出关系 (95a, b)，因为我们有

$$\nabla\times(\nabla s)=(\nabla\times\nabla)s=0,$$

---

1) 在旋度的情形，这个陈述必须说明一下. 一般说来，如同在第 195 页说过的，两个向量的向量积的大小和方向有一定的几何意义、但有一个例外的事情，就是当我们改变所用笛卡儿坐标系的定向时，向量积就要变到相反的位置. 这意味着，只要我们不改变坐标系的定向（也就是只要我们仅仅使用行列式为 +1 的正交变换），对于一个向量 $\mathbf{U}$, $\operatorname{curl}\mathbf{U}=\nabla\times\mathbf{U}$ 总是一个固定的向量；改变坐标系定向的结果是 $\operatorname{curl}\mathbf{U}$ 变到它的相反位置.

$$\nabla\cdot(\nabla\times\mathbf{U})=\det(\nabla,\nabla,\mathbf{U})=0.$$

另外一个十分重要的向量微分算子的组合是梯度的散度

$$\begin{aligned}\operatorname{div}\operatorname{grad}s&=\nabla\cdot(\nabla s)\\&=\frac{\partial^2 s}{\partial x_1^2}+\frac{\partial^2 s}{\partial x_2^2}+\frac{\partial^2 s}{\partial x_3^2}=\Delta s.\end{aligned}\tag{95c}$$

这里

$$\Delta=\nabla\cdot\nabla=\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial x_3^2}\tag{95d}$$

是熟知的“拉普拉斯算子 (Laplace operator)”. 偏微分方程

$$\Delta s=\frac{\partial^2 s}{\partial x_1^2}+\frac{\partial^2 s}{\partial x_2^2}+\frac{\partial^2 s}{\partial x_3^2}=0\tag{95e}$$

被数学物理中很多重要的数量场 s 所满足，被称作“拉普拉斯方程”或“位势 (potential) 方程”.

当独立变量的个数不是三个时，“向量分析”这个术语也常常使用. 由 n 个独立变量 $x_1,x_2,\cdots,x_n$ 的 n 个函数 $u_1,u_2,\cdots,u_n$ 所形成的一个函数组确定一个 n 维空间中的一个向量场. 这里关于一个数量场的梯度的概念和拉普拉斯算子的概念仍保持着它们的意义. 类似于向量场旋度的概念变得更复杂了. 在 n 维空间中，类似于 (95a, b) 的关系，最满意的研究途径是通过外微分型的演算，这将在下一章中讨论.

d. 向量族. 在空间曲线论和质点运动中的应用

除向量场外，我们还考虑单个参量的向量流型，称作向量族，其中向量 $\mathbf{U}=(u_1,u_2,u_3)$ 不是对应着一个空间区域的每个一点，而是对应着单个参量 t 的每一个值. 我们写作 $\mathbf{U}=\mathbf{U}(t)$. 向量 $\mathbf{U}$ 的微商可自然地定义作

$$\frac{d\mathbf{U}}{dt}=\lim_{h\to 0}\frac{1}{h}[\mathbf{U}(t+h)-\mathbf{U}(t)],\tag{96a}$$

它的分量显然是

$$\frac{du_1}{dt}, \frac{du_2}{dt}, \frac{du_3}{dt}. \tag{96b}$$

容易证明，这种向量微分法满足类似于通常的微商规则

$$\frac{d}{dt}(\mathbf{U}+\mathbf{V}) = \frac{d}{dt}\mathbf{U} + \frac{d}{dt}\mathbf{V}; \tag{97a}$$

$$\frac{d}{dt}(\lambda\mathbf{U}) = \frac{d\lambda}{dt}\mathbf{U} + \lambda\frac{d}{dt}\mathbf{U},$$

$$\frac{d}{dt}(\mathbf{U}\cdot\mathbf{V}) = \mathbf{U}\cdot\frac{d\mathbf{V}}{dt} + \frac{d\mathbf{U}}{dt}\cdot\mathbf{V}, \tag{97b}$$

$$\frac{d}{dt}(\mathbf{U}\times\mathbf{V}) = \mathbf{U}\times\frac{d\mathbf{V}}{dt} + \frac{d\mathbf{U}}{dt}\times\mathbf{V}. \tag{97c}$$

我们能把这些概念应用到由参量表示式

$$x_1 = \phi_1(t), \quad x_2 = \phi_2(t), \quad x_3 = \phi_3(t)$$

所给出的空间中一条曲线上的点 P 的 **位置向量** $\mathbf{X} = \mathbf{X}(t) = \overrightarrow{OP}$ 所组成的向量族. 于是

$$\mathbf{X} = (x_1, x_2, x_3) = (\phi_1(t), \phi_2(t), \phi_3(t)).$$

向量 $\dfrac{d\mathbf{X}}{dt}$ 的方向是曲线在对应于 t 的点的 **切线** 的方向. 因为向量 $\Delta\mathbf{X} = \mathbf{X}(t+\Delta t) - \mathbf{X}(t)$ 的方向是连接参量值 t 和 $t+\Delta t$ 所对应的点的线段的方向. 这在 $\Delta t > 0$ 时，对于向量 $\dfrac{\Delta\mathbf{X}}{\Delta t}$ 同样成立. 当 $\Delta t \to 0$, 这个弦的方向就趋于切线的方向. 如果我们引进一个新参量，即从一个确定的起点测得的曲线的弧长 s, 并用 s 代替 t, 我们就能证明

$$\left|\frac{d\mathbf{X}}{ds}\right|^2 = \frac{d\mathbf{X}}{ds}\cdot\frac{d\mathbf{X}}{ds} = 1. \tag{98}$$

证明的思路是同关于平面曲线的相应的证明 (第一卷第 399 页) 完全一样的. 因此, $\dfrac{d\mathbf{X}}{ds}$ 是一个单位向量. 将方程 (98) 两边对 s 微分，应用 (97b), 就得到

$$\frac{d\mathbf{X}}{ds}\cdot\frac{d^2\mathbf{X}}{ds^2} + \frac{d^2\mathbf{X}}{ds^2}\cdot\frac{d\mathbf{X}}{ds} = 2\frac{d\mathbf{X}}{ds}\cdot\frac{d^2\mathbf{X}}{ds^2} = 0. \tag{99}$$

这个方程表明向量

$$\frac{d^2\mathbf{X}}{ds^2}=\left(\frac{d^2x_1}{ds^2},\frac{d^2x_2}{ds^2},\frac{d^2x_3}{ds^2}\right)$$

垂直于切线. 我们称这个向量为 **曲率向量** 或 **主法向量**, 并且称它的长度

$$\kappa=\frac{1}{\rho}=\left|\frac{d^2\mathbf{X}}{ds^2}\right| \tag{100}$$

为曲线在该对应点处的 **曲率**. 和过去一样, 称它的倒数 $\rho=\dfrac{1}{\kappa}$ 为 **曲率半径**. 从曲线上的这个点沿主法向量的方向测量一个距离为 ρ 的点, 称为 **曲率中心**.

我们将证明曲率的这个定义与第一卷 (第 399 页) 中给出的平面曲线的曲率是一致的. 对于每一个 s, $\mathbf{Y}=\dfrac{d\mathbf{X}}{ds}$ 是长度为 1 并且方向为切线方向的向量. 如果我们设想 $\mathbf{Y}(s+\Delta s)$ 和 $\mathbf{Y}(s)$ 是以原点为公共起点的向量, 则差 $\Delta\mathbf{Y}=\mathbf{Y}(s+\Delta s)-\mathbf{Y}(s)$ 是连接它们末点的向量. 曲线上对应于参数 s 和 $s+\Delta s$ 的点的切线之间的夹角 β 等于向量 $\mathbf{Y}(s)$ 与 $\mathbf{Y}(s+\Delta s)$ 之间的夹角. 于是

$$|\Delta\mathbf{Y}|=|\mathbf{Y}(s+\Delta s)-\mathbf{Y}(s)|=2\sin\frac{\beta}{2},$$

因为

$$|\mathbf{Y}(s)|=|\mathbf{Y}(s+\Delta s)|=1.$$

利用

$$\frac{2\sin\beta/2}{\beta}\to 1,\ \text{当}\ \beta\to 0,$$

就得到

$$\left|\frac{d^2\mathbf{X}}{ds^2}\right|=\left|\frac{d\mathbf{Y}}{ds}\right|=\lim_{\Delta s\to 0}\left|\frac{\Delta Y}{\Delta s}\right|=\lim_{\Delta s\to 0}\frac{\beta}{\Delta s}.$$

因此, κ **是曲线上两点切线之间的夹角与这两点之间弧长的比值, 当这两点互相趋近时的极限**. 而这个极限确定了平面曲线的

曲率[1].

曲率向量在力学中起着重要的作用. 我们假定一个质点沿一条曲线运动, 在时刻 t 有位置向量 $\mathbf{X}(t)$. 则运动速度的大小和方向二者都由向量 $\frac{d\mathbf{X}}{dt}$ 给出. 类似地, 加速度由向量 $\frac{d^2\mathbf{X}}{dt^2}$ 给出. 由链式法则, 我们有

$$\frac{d\mathbf{X}}{dt}=\frac{ds}{dt}\frac{d\mathbf{X}}{ds}$$

和

$$\frac{d^2\mathbf{X}}{dt^2}=\frac{d^2s}{dt^2}\frac{d\mathbf{X}}{ds}+\left(\frac{ds}{dt}\right)^2\frac{d^2\mathbf{X}}{ds^2}. \tag{101}$$

按我们已经知道的向量 $\mathbf{X}$ 关于 s 的一阶和二阶导数, 方程 (101) 表示了如下事实: 运动的 **加速度向量** 是两个向量的和. 它们中的一个是沿曲线的切线方向, 它的长度等于 $\frac{d^2s}{dt^2}$, 即点在它的轨道方向上的加速度 (速度变化率或 **切线加速度**). 另一个的方向是垂直于轨道的方向且指向曲率中心, 它的长度等于速度的平方与曲率的乘积 (这就是 **法线加速度**). 对于一个单位质量的质点, 加速度向量就等于作用于质点上的力. 如果没有力作用在曲线的方向上 (如在这种情形, 一个质点受到约束而沿着一条曲线运动, 仅仅受到垂直于曲线的反作用力), 切线加速度就等于零, 并且整个加速度垂直于曲线, 大小为速度的平方乘以曲率.

练　习　2.5

1. 验证一点 Q 相对于一点 P 的位置向量 $\overrightarrow{PQ}$ 在坐标变换下也是个向量.

2. 推导下列恒等式:

---

1) 在空间曲线的情形, 我们不能像平面曲线那样, 把 β 看成与倾角 α 的增量 $\Delta\alpha$ 恒等. 理由是 $\mathbf{Y}(s)$ 与 $\mathbf{Y}(s+\Delta s)$ 之间的夹角一般不等于向量 $\mathbf{Y}(s)$ 与 $\mathbf{Y}(s+\Delta t)^s$ 同某一固定的第三方向之间的夹角的差. **空间中方向之间的夹角, 不像平面上的一样, 它不是可加的**.

(a) $\text{grad}\,(\alpha\beta) = \alpha\text{grad}\,\beta + \beta\text{grad}\,(\alpha)$,

(b) $\text{div}\,(\alpha\mathbf{U}) = \mathbf{U}\cdot\text{grad}\,\alpha + \alpha\text{div}\,\mathbf{U}$,

(c) $\text{curl}\,(\alpha\mathbf{U}) = \text{grad}\,\alpha\times\mathbf{U} + \alpha\text{curl}\,\mathbf{U}$,

(d) $\text{div}\,(\mathbf{U}\times\mathbf{V}) = \mathbf{V}\cdot\text{curl}\,\mathbf{U} - \mathbf{U}\cdot\text{curl}\,\mathbf{V}$.

3. 设 $\mathbf{U}\cdot\nabla$ 是算子

$$\mathbf{U}_x\frac{\partial}{\partial x} + \mathbf{U}_y\frac{\partial}{\partial y} + \mathbf{U}_z\frac{\partial}{\partial z}$$

的符号，验证

(a) $\text{grad}\,(\mathbf{U}\cdot\mathbf{V}) = \mathbf{U}\cdot\nabla\mathbf{V} + \mathbf{V}\cdot\nabla\mathbf{U} + \mathbf{U}\times\text{curl}\,\mathbf{V} + \mathbf{V}\times\text{curl}\,\mathbf{U}$,

(b) $\text{curl}\,(\mathbf{U}\times\mathbf{V}) = \mathbf{U}\text{div}\,\mathbf{V} - \mathbf{V}\text{div}\,\mathbf{U} + \mathbf{V}\cdot\nabla\mathbf{U} - \mathbf{U}\cdot\nabla\mathbf{V}$.

4. 对于拉普拉斯算子 Δ 建立

$$\Delta\mathbf{U} = \text{grad}\,\text{div}\,\mathbf{U} - \text{curl}\,\text{curl}\,\mathbf{U}.$$

5. 试求曲线 $x = f(t), y = g(t), z = h(t)$ 在点 $t = t_0$ 的所谓“密切平面”的方程，即经过曲线的三个点的平面，当这些点趋向于参数 t_0 所对应的点时的极限.

6. 证明曲率向量和切线向量二者都位于密切平面内.

7. 设 C 是具有连续转动的切线的一条光滑曲线. 设 d 为曲线上两点间的最短距离，又设 l 为这两点间的弧长. 证明，当 d 很小时有 $d - l = o(d)$.

8. 设 t 为任意参量，证明曲线 $\mathbf{X} = \mathbf{X}(t)$ 的曲率是

$$k = \frac{\{|\mathbf{X}'|^2|\mathbf{X}''|^2 - (\mathbf{X}'\cdot\mathbf{X}'')^2\}^{\frac{1}{2}}}{|\mathbf{X}'|^3}.$$

9. 如果 $\mathbf{X} = \mathbf{X}(t)$ 是一条曲线的任何一个参数表示式，则以 $\mathbf{X}$ 为起点的向量 $d^2\mathbf{X}/dt^2$ 在 $\mathbf{X}$ 点的密切平面内.

10. 如果 C 是一条连续可微的闭曲线， A 为不在 C 上的一点，则在 C 上有一点 B，使得从 A 到 B 的距离比较 A 到 C 上其他点的距离都要短. 证明 AB 垂直于曲线.

11. 在柱面 $x^2+y^2=a^2$ 上，画一条曲线，使得这曲线上任何一点 P 处的切线与 z 轴之间的夹角等于 y 轴与柱面在 P 点处的切平面之间的夹角. 证明曲线上任何一点 P 的坐标能够用一个参量 θ 通过以下方程组来表示：

$$x=a\cos\theta,\quad y=a\sin\theta,\quad z=C\pm\log\sin\theta$$

再证明这条曲线的曲率是 $\dfrac{1}{a}\sin\theta(1+\sin^2\theta)^{\frac{1}{2}}$.

12. 求曲线 $x=\cos\theta$, $y=\sin\theta$, $z=f(\theta)$ 在点 θ 的密切平面的方程 (参看第 5 题). 证明，如果 $f(\theta)=(\cosh A\theta)/A$, 则每一个密切平面都与中心为原点、半径为 $\sqrt{(1+1/A^2)}$ 的球面相切.

13. (a) 证明经过曲线

$$x=\frac{1}{3}at^3,\quad y=\frac{1}{2}bt^2,\quad z=ct$$

上的三点 t_1,t_2,t_3 的平面具有方程

$$\frac{3x}{a}-2(t_1+t_2+t_3)\frac{y}{b}+(t_2t_3+t_3t_1+t_1t_2)\frac{z}{c}-t_1t_2t_3=0.$$

(b) 证明在 t_1,t_2,t_3 处的密切平面的交点在这个平面上.

14. 设 $\mathbf{X}=\mathbf{X}(s)$ 是任意一条空间曲线，使得向量 $\mathbf{X}(s)$ 三次连续可微 (s 为弧长). 求在点 s 处与曲线最紧密切触的球的中心.

15. 如果 $\mathbf{X}=\mathbf{X}(s)$ 是单位球面上一条曲线，其中 s 为弧长，则

$$|\dddot{\mathbf{X}}|^2-|\ddot{\mathbf{X}}|^4=|\dddot{\mathbf{X}}|^2-(\dot{\mathbf{X}}\cdot\dddot{\mathbf{X}})^2=(\ddot{\mathbf{X}}\cdot[\dot{\mathbf{X}}\times\dddot{\mathbf{X}}])^2.$$

16. 一条曲线上两个邻近点的密切平面之间的夹角与该两点间弧长之比的极限 (即单位法向量相对于弧长的导数) 称作这曲线的**挠率**. 设 $\xi_1(s)$、$\xi_2(s)$ 表示沿着曲线 $\mathbf{X}(s)$ 切线方向和曲率向量方向的单位向量；我们用 $\xi_3(s)$ 表示垂直于 ξ_1 和 ξ_2 的单位向量 (所谓**副法向量**)，它由 $[\xi_1\times\xi_2]$ 给出. 证明弗雷乃公式

$$\dot{\xi}_1=\frac{\xi_2}{\rho},$$

$$\dot{\xi}_2 = -\frac{\xi_1}{\rho} + \frac{\xi_3}{\tau},$$
$$\dot{\xi}_3 = -\frac{\xi_2}{\tau},$$

其中 $1/\rho = k$ 是 $x(s)$ 的曲率， $1/\tau$ 是 $\mathbf{X}(s)$ 的挠率.

17. 使用练习 16 中的向量 ξ_1, ξ_2, ξ_3 作为坐标向量，求 (a) 向量 $\dddot{\mathbf{X}}$ 的表达式， (b) 从点 $\mathbf{X}$ 到与 $\mathbf{X}$ 最紧密切触的球的中心的向量的表达式.

18. 证明挠率为零的曲线是平面曲线.

19. 考虑空间一个定点 A 和一个动点 P, 它的运动由时间的一个函数给出. 用 $\dot{P}$ 表示点 P 的速度向量，且用 $\mathbf{a}$ 表示 P 到 A 方向的单位向量，证明

$$\frac{d}{dt}|\overrightarrow{PA}| = -\mathbf{a} \cdot \dot{\mathbf{P}}.$$

20. (a) 设 A, B, C 是三个固定的不共线的点，设 P 为一个动点. 设 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 分别为 PA, PB, PC 方向的单位向量；把速度向量 $\dot{\mathbf{P}}$ 表示成这些向量的线性组合：

$$\dot{\mathbf{P}} = \mathbf{a}u + \mathbf{b}v + \mathbf{c}w.$$

证明

$$\dot{\mathbf{a}} = \frac{1}{|A-P|}\{[(\mathbf{a}\cdot\mathbf{b})v + (\mathbf{a}\cdot\mathbf{c})w]\mathbf{a} - v\mathbf{b} - w\mathbf{c}\}.$$

(b) 证明点 P 的加速度向量 $\ddot{\mathbf{P}}$ 是

$$\ddot{\mathbf{P}} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c},$$

其中

$$\begin{aligned}\alpha =& \dot{u} + uv\left(\frac{\mathbf{a}\cdot\mathbf{b}}{|A-P|} - \frac{1}{|B-P|}\right)\\ &+ uw\left(\frac{\mathbf{a}\cdot\mathbf{c}}{|A-P|} - \frac{1}{|C-P|}\right),\end{aligned}$$

以及 β 和 γ 具有类似的表达式.

21. 设 $z = u(x,y)$ 表示由任意一条曲线的切线形成的曲面. 证明 (a) 这曲线的每一个密切平面都是该曲面的切平面, (b) 函数 $u(x,y)$ 满足方程

$$u_{xx}u_{yy} - u_{xy}^2 = 0.$$

第三章　微分学的发展和应用

3.1　隐 函 数

a. 一般说明

在解析几何中一条曲线的方程常常不是用 $y = f(x)$ 这个形式而是用 $F(x,y) = 0$ 这个形式给出的；譬如一条直线就由方程 $ax+by+c = 0$ 表示，而一个椭圆就由方程 $x^2/a^2+y^2/b^2 = 1$ 表示. 要得到曲线的形如 $y = f(x)$ 的方程，我们就必须从方程 $F(x,y) = 0$ 把 y 解出来. 在第一卷中我们曾考虑过这样一个特殊的问题，就是找 $y = f(x)$ 的反函数，这个问题实际上就是从方程 $F(x,y) = y - f(x) = 0$ 解出 x 的问题.

这些例子揭示了从方程 $F(x,y) = 0$ 解出 x 或 y 的方法的重要性. 甚至对包含多于两个变量的函数的方程我们也将找到这种方法.

在最简单的情形，例如前述直线与椭圆的方程，方程的解可以马上表为初等函数的形式. 而在另外的情形，解可以被近似，并且要怎样近似就可以怎样近似. 不过，对于很多目的来说，宁可不去求方程的解出的形式，或是它们的近似，而是要由直接研究函数 $F(x,y)$ 来得到关于方程的解的一些结论. 这里，没有假定 $F(x,y)$ 中的变量 x,y 那一个比另一个处于优先地位.

并不是每一个方程 $F(x,y) = 0$ 都是某个函数 $y = f(x)$ 或 $x = \phi(y)$ 的隐式表示. 很容易给出方程 $F(x,y) = 0$ 的例子，它不

允许任何一元函数作为它的解. 例如: 方程 $x^2+y^2=0$ 只被单独的一对值 $x=0, y=0$ 所满足, 而方程 $x^2+y^2+1=0$ 则根本不被任何实数值所满足. 这就是为什么有必要更严密地考察一个方程 $F(x,y)=0$ 能够定义一个函数 $y=f(x)$ 的条件以及这个函数能具有的性质的原因.

练　习　3.1 a

1. 设对于若干对值 (a,b) 有 $f(a,b)=0$. 如 a 为已知, 请给出一个寻找 b 的构造性的迭代方法. 问 f 要具有什么条件, 这个方法才可行?

b. 几何解释

为了使情况直观, 我们把函数 $F(x,y)$ 用三维空间的曲面 $z=F(x,y)$ 来表示. 方程 $F(x,y)=0$ 的解完全等同于两个方程 $z=F(x,y)$ 与 $z=0$ 的联立解. 从几何上看, 我们的问题就是要确定曲面 $z=F(x,y)$ 是否与 xy 坐标平面相交于某条曲线 $y=f(x)$ 或 $x=\phi(y)$. (至于这交线能够延伸多远, 我们这里并不管它.)

第一个可能性是曲面与平面没有公共点. 例如抛物面 $z=F(x,y)=x^2+y^2+1$ 完全位于 xy 平面之上, 因而这里没有任何交线. 很清楚, 我们仅需考虑最少有一个点 (x_0,y_0) 使 $F(x_0,y_0)=0$ 的情形. 这个点 (x_0,y_0) 构成了我们解的 "初始点".

知道了一个初始解, 接着就有两个可能: 在点 (x_0,y_0) 处的切平面, 或者是水平的, 或不是水平的. 如果切平面是水平的, 则我们马上可以用例子表明, 不可能从 (x_0,y_0) 扩展成一个解 $y=f(x)$ 或 $x=\phi(y)$. 例如, 抛物面 $z=x^2+y^2$ 有初始解 $x=0, y=0$, 但是它不再包含 xy 平面上的其他点. 与此相对比, 曲面 $z=xy$ 有初始解 $x=0, y=0$, 且与 xy 平面沿着直线 $x=0$ 与直线 $y=0$ 相交, 但是却不存在原点的任何邻域, 使得我们能够对**整个**交线用一个函数 $y=f(x)$ 或一个函数 $x=\phi(y)$ 表示出来 (参看图 3.1 与

3.2). 另一方面，甚至当初始解的切平面是水平时，$F(x,y)=0$ 也很可能有一个解，如像 $F(x,y)=(y-x)^4=0$ 的情况．所以在水平切平面这个特殊情况，没有可能作出一个一般确定的结论．

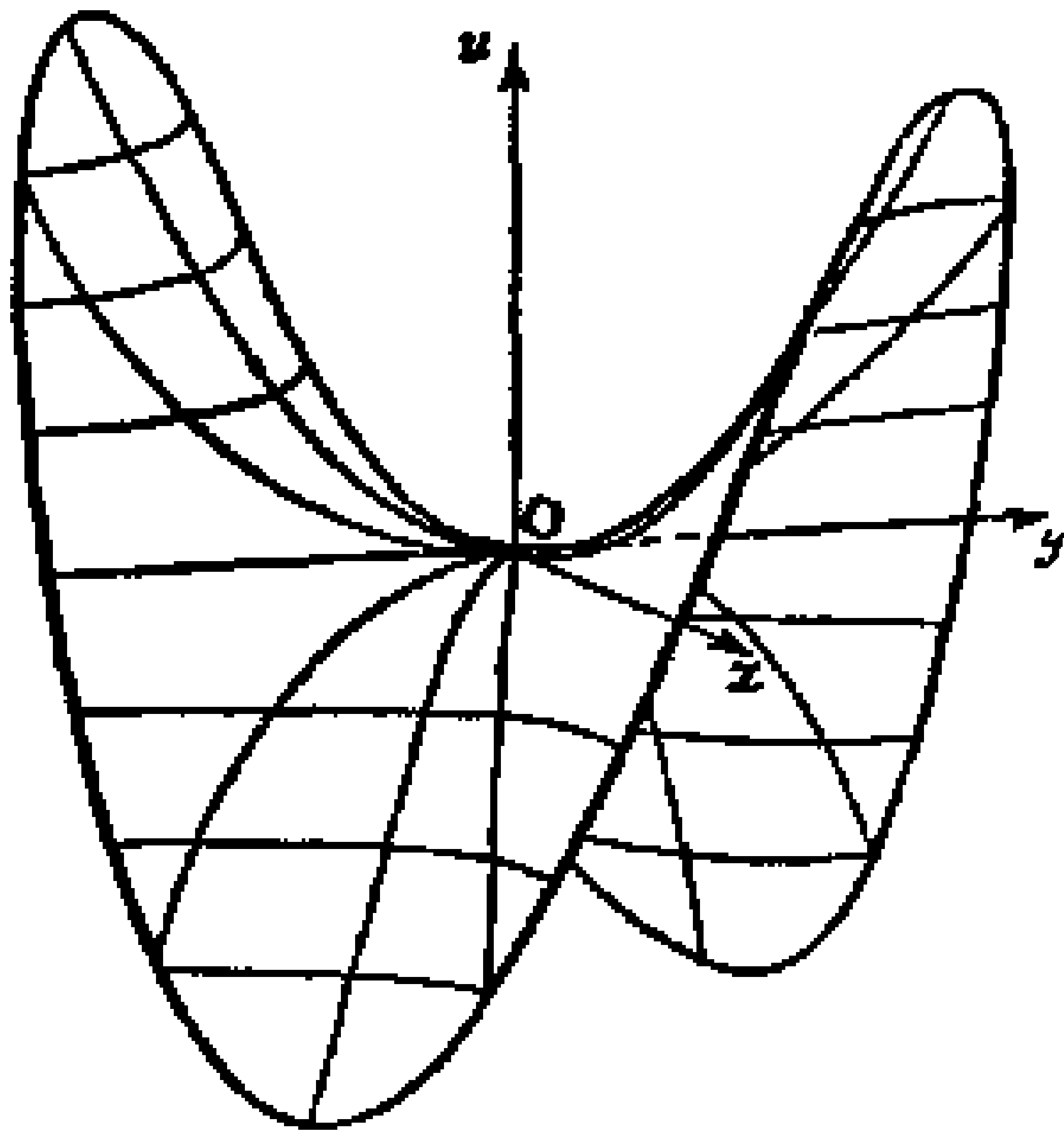


图 3.1　曲面 $u=xy$

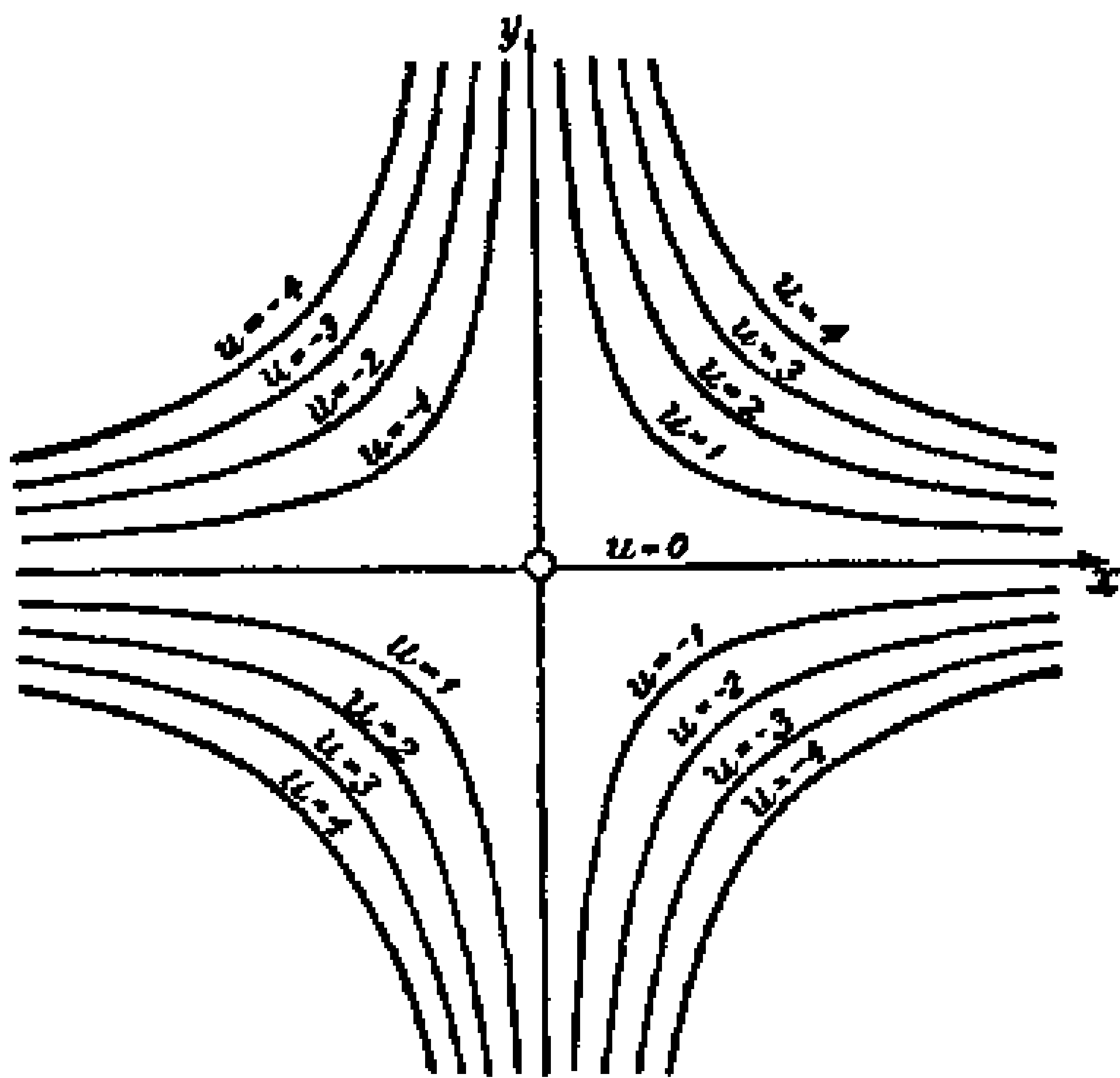


图 3.2　$u=xy$ 的等高线

剩下的可能性便是初始点处的切平面不是水平的．这时，直观地想到曲面 $z=F(x,y)$ 可以在原始解的邻域内由切平面来近似，

我们就可以期望曲面尽可能弯得贴切一些，使得不可避免地要在 (x_0, y_0) 附近交 xy 平面于一个单独的完全确定的交线，并且期望这一曲线邻近初始解的一部分可以由等式 $y = f(x)$ 或 $x = \phi(y)$ 来表示. 从解析式来看，切平面不是水平意味着 $F_x(x_0, y_0)$ 和 $F_y(x_0, y_0)$ 总有一个不是 0(见第 53 页). 这些就是下一节讨论的基础.

练 习 3.1 b

1. 研究曲面 $z = f(x, y)$，确定在指定点 (x_0, y_0) 的一个邻域内能否把方程 $f(x, y) = 0$ 中的 y 解出成为 x 的一个函数：

(a) $f(x, y) = x^2 - y^2$, $x_0 = y_0 = 0$.

(b) $f(x, y) = [\log(x + y)]^{\frac{1}{2}}$, $x_0 = 1.5$, $y_0 = -0.5$.

(c) $f(x, y) = \sin[\pi(x + y)] - 1$, $x_0 = y_0 = 1/4$.

(d) $f(x, y) = x^2 + y^2 - y$, $x_0 = y_0 = 0$.

c. 隐函数定理

我们现在叙述隐函数存在的充分条件，并且同时给出微分它们的一个法则.

设 $F(x, y)$ 在点 (x_0, y_0) 的一个邻域内有连续的偏微商 F_x 和 F_y，其中

$$F(x_0, y_0) = 0, \quad F_y(x_0, y_0) \neq 0. \tag{1}$$

则以点 (x_0, y_0) 为中心，存在某个矩形.

$$x_0 - a \leq x \leq x_0 + a, \quad y_0 - \beta \leq y \leq y_0 + \beta \tag{2}$$

使得对 $x_0 - a \leq x \leq x_0 + a$ 给出的区间 I 上的每一个 x，方程

$$F(x, y) = 0$$

恰好有一个解 $y = f(x)$ 位于区间 $y_0 - \beta \leq y \leq y_0 + \beta$ 上. 这个函

数 f **满足初条件** $y_0 = f(x_0)$, **并且对** I **内的每一个** x **都有**

$$F(x, f(x)) = 0. \tag{3}$$

$$y_0 - \beta \leq f(x) \leq y_0 + \beta, \tag{3a}$$

$$F_y(x, f(x)) \neq 0. \tag{3b}$$

此外, f **是连续的**, **并在** I **上有由方程**

$$y' = f'(x) = -\frac{F_x}{F_y} \tag{4}$$

给出的连续微商.

这是一个在初始解 (x_0, y_0) 的邻域内来解方程 $F(x, y) = 0$ 的严格的 **局部性** 存在定理. 这个定理并没有指出如何去找"初始解", 或是如何确定方程 $F(x, y) = 0$ 是否有满足它的任何一个解 (x, y). 这些是 **整体性** 的问题, 已经超出了定理范围. 解 $y = f(x)$ 的 **唯一性** 与 **正则性** 也只能是局部地保证成立, 即当 y 限制在区间 $y_0 - \beta < y < y_0 + \beta$ 之内的时候. 这种限制的必要性可从方程

$$F(x, y) = x^2 + y^2 - 1 = 0$$

这样简单的例子明显地看出. 对每一个满足 $-1 < x < 1$ 的 x, 这方程有两支不同的解

$$y = \pm\sqrt{1 - x^2}.$$

对每一个 x 任意选定一个符号, 就得到一个单值解 $y = f(x)$. 很清楚, 由这种方法我们能够得到每一个 x 都不连续的解; 譬如对有理的 x 取正号, 而对无理的 x 取负号. 当我们限定 y 取不变的符号时才得到连续解. 这个符号可以这样选定, 即对给定的 x_0, $-1 < x_0 < 1$, 在满足 $x_0^2 + y_0^2 = 1$ 的两个可能的值 y_0 中选定一个. 一个满足 $y_0 = f(x_0)$ 的唯一的连续解就可以对所有 x, $-1 < x < 1$,

取既满足 $x^2+y^2=1$ 又与 y_0 同号的 y 而得到. 从几何上说, f 的图像是包含点 (x_0,y_0) 的那个上半圆或下半圆. 函数 f 有连续的微商

$$y'=-\frac{F_x}{F_y}=-\frac{x}{y}=-\frac{x}{f(x)},\quad -1<x<1.$$

对于 $x=\pm1$ 定义 y 为 0, 这样解 $y=f(x)$ 就在闭区间 $-1\le x\le 1$ 上连续. 不过在区间的端点, 由于 $F_y=0$ 微商 y' 变成无穷大.

下一节我们将对这个一般性的定理给以证明. 这里我们看到, 只要满足 (3) 的函数 $f(x)$ 的存在性与可微性一旦成立, 我们就可以应用链式法则 (见第 61 页 (18)) 微分 $F(x,y)$ 来得到 $f'(x)$ 的显式表示. 由此推出

$$F_x+F_y\cdot f'(x)=0,$$

因而只要 $F_y\neq0$, 就导出公式 (4). 等价地, 如果方程 $F(x,y)=0$ 确定 y 作为 x 的一个函数, 我们就有

$$dF=F_xdx+F_ydy=0,$$

因此

$$dy=\frac{dy}{dx}dx=-\frac{F_x}{F_y}dx.$$

一个隐函数 $y=f(x)$ 可以被微分到任意给定阶, 假定函数 $F(x,y)$ 具有那同样阶的连续偏微商. 举例来说, 如果 $F(x,y)$ 在矩形 (2) 中具有连续的一阶与二阶的微商, 方程 (4) 的右边便是 x 的一个复合函数

$$-\frac{F_x(x,f(x))}{F_y(x,f(x))}.$$

因为, 由于 (3b) 分母不为 0, 又由于已知 $f(x)$ 具有一阶的连续微商, 我们就由 (4) 断定 y' 有连续微商; 再应用链式法则就得到

$$y''=-\frac{F_yF_{xx}+F_yF_{xy}f'-F_xF_{yx}-F_xF_{yy}f'}{F_y^2}.$$

对 f' 用表达式 (4) 代入后，我们得到

$$y'' = -\frac{F_y^2 F_{xx} - 2F_{xy}F_xF_y + F_x^2 F_{yy}}{F_y^3}. \tag{5}$$

当我们应用隐函数的一般定理肯定了某区间内函数 f 的存在性后，法则 (4) 和 (5) 便可用来求出隐函数 $y = f(x)$ 的微商，即使不可能把 y 明显表为初等函数 (有理函数，三角函数等). 另外，即使我们可把方程 $F(x, y) = 0$ 按 y 显式解出，由法则 (4) 和 (5) 来找出 y 的微商也常常比较方便，而不用去求那 $y = f(x)$ 的任何显式表示.

例

1. 双纽线 (第一卷第 113 页) 方程

$$F(x,y) = (x^2 + y^2)^2 - 2a^2(x^2 - y^2) = 0$$

要对 y 解出来是不容易的. 对于 $x = 0, y = 0$, 我们求得

$$F = 0, \quad F_x = 0, \quad F_y = 0.$$

我们的定理在此无效，正如我们由这一事实就可料到，即通过原点有双纽线的两个不同分支. 不过，对曲线上所有 $y \neq 0$ 的点我们的法则都可以应用，因而函数 $y = f(x)$ 的微商由下式给出

$$y' = -\frac{F_x}{F_y} = -\frac{4x(x^2 + y^2) - 4a^2x}{4y(x^2 + y^2) + 4a^2y}.$$

我们可以从方程本身得到关于曲线的许多重要知识，而用不着 y 的显式表示. 例如，极大与极小可以出现在 $y' = 0$ 处，即当 $x = 0$, 或是 $x^2 + y^2 = a^2$ 时. 由双纽线的方程知道，当 $x = 0$ 时 $y = 0$; 但是在原点却没有极值点 (参看第一卷第 113 页图 S. 1.3). 所以，这两个方程给出了四个点

$$\left(\pm\frac{a}{2}\sqrt{3}, \quad \pm\frac{a}{2}\right)$$

作为极大和极小.

2. 笛卡儿叶形线

具有方程

$$F(x,y)=x^3+y^3-3axy=0$$

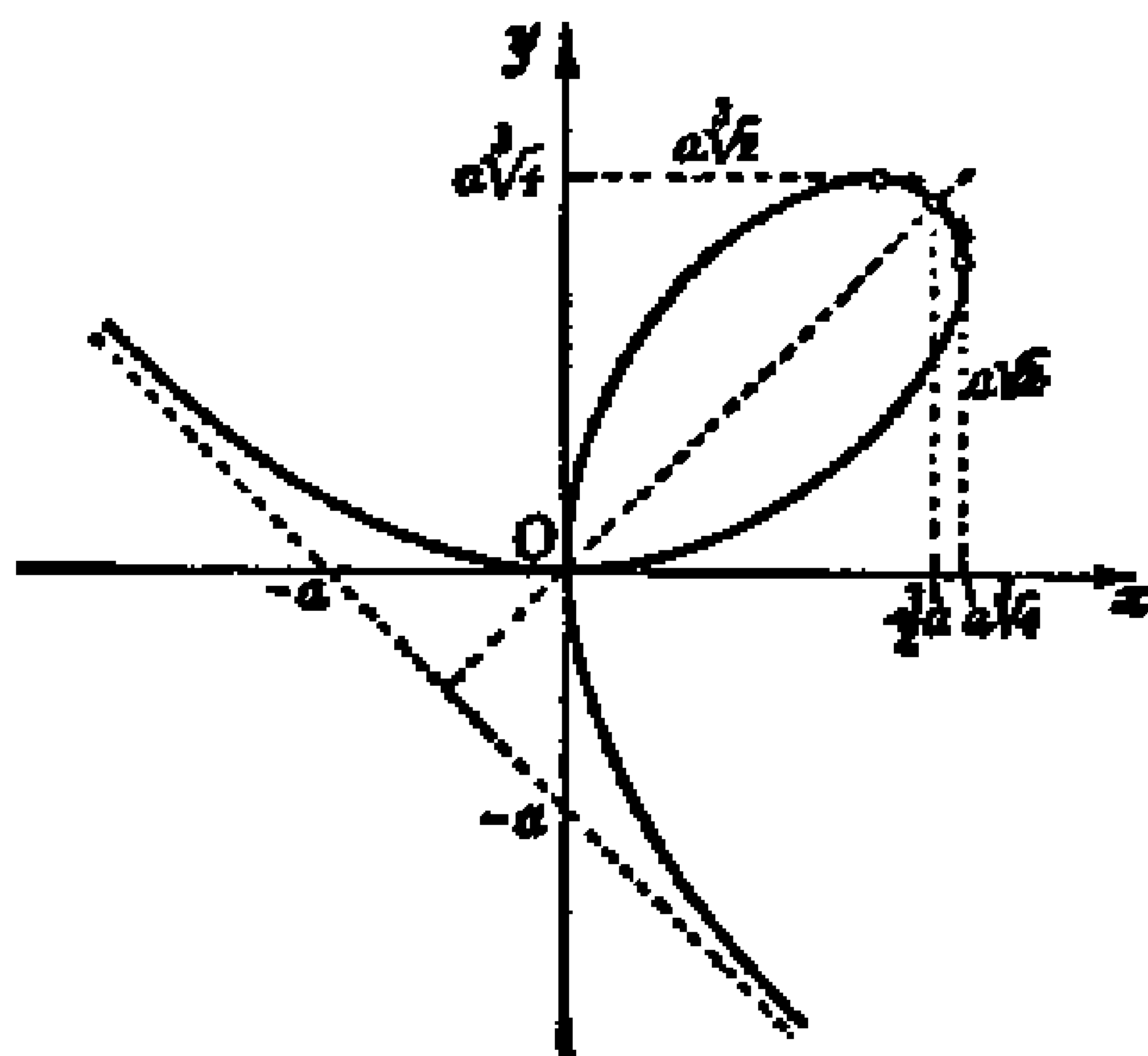


图 3.3　笛卡儿叶形线

(参看图 3.3), 其显式解很难看. 在原点处曲线与自己相交, 由于在该点 $F=F_x=F_y=0$, 我们的法则又一次失效. 对于所有使 $y^2\neq ax$ 的点, 我们有

$$y'=-\frac{F_x}{F_y}=-\frac{x^2-ay}{y^2-ax}.$$

因此, 当 $x^2=ay$ 时, 或者按照曲线的方程, 即当

$$x=a\sqrt[3]{2},\quad y=a\sqrt[3]{4}$$

时, 微商有零点.

练　习　3.1 c

1. 证明　下列各方程在指定点附近有对 y 的唯一解:

(a) $x^2+xy+y^2=7$ $(2,1)$,

(b) $x\cos xy=0$ $\left(1,\dfrac{\pi}{2}\right)$,

(c) $xy + \log xy = 1\ (1,1)$,

(d) $x^5 + y^5 + xy = 3(1,1)$.

2. 求出习题 1 中的各个解的一阶微商，并在指定点处给出它们的值.

3. 求出习题 1 中各个解的二阶微商，并在指定点处给出它们的值.

4. 习题 1 中的隐式定义的函数中，哪一个在指定点是凸的？

5. 找出满足方程 $x^2 + xy + y^2 = 27$ 的函数 y 的极大值与极小值.

6. 设 $f_y(x,y)$ 在点 (x_0, y_0) 的一个邻域内是连续的. 试证方程

$$y = y_0 + \int_{x_0}^{x} f(\xi, y) d\xi$$

在围绕 $x = x_0$ 的某区间内把 y 确定为 x 的函数.

d. 隐函数定理的证明

隐函数的存在性系由中值定理直接推出 (见第一卷第 48 页).

假设 $F(x,y)$ 在点 (x_0, y_0) 的一个邻域内有定义并有连续的一阶微商，并令

$$F(x_0, y_0) = 0, \quad F_y(x_0, y_0) \neq 0.$$

不失一般性，我们假设 $m = F_y(x_0, y_0) > 0$. 否则，我们只要用 $-F$ 来代替 F 就行了，这并不改变由方程 $F(x,y) = 0$ 所描绘的点集. 因为 $F_y(x,y)$ 是连续的，我们可以找到一个以 (x_0, y_0) 为中心的矩形 R, 小到整个 R 可以完全落在 F 的定义域之内，而在整个 R 内 $F_y(x,y) > m/2$. 设 R 是矩形

$$x_0 - a \le x \le x_0 + a, \quad y_0 - \beta \le y \le y_0 + \beta$$

(见图 3.4). 由于 $F_x(x,y)$ 也是连续的，我们推断 F_x 在 R 是有界

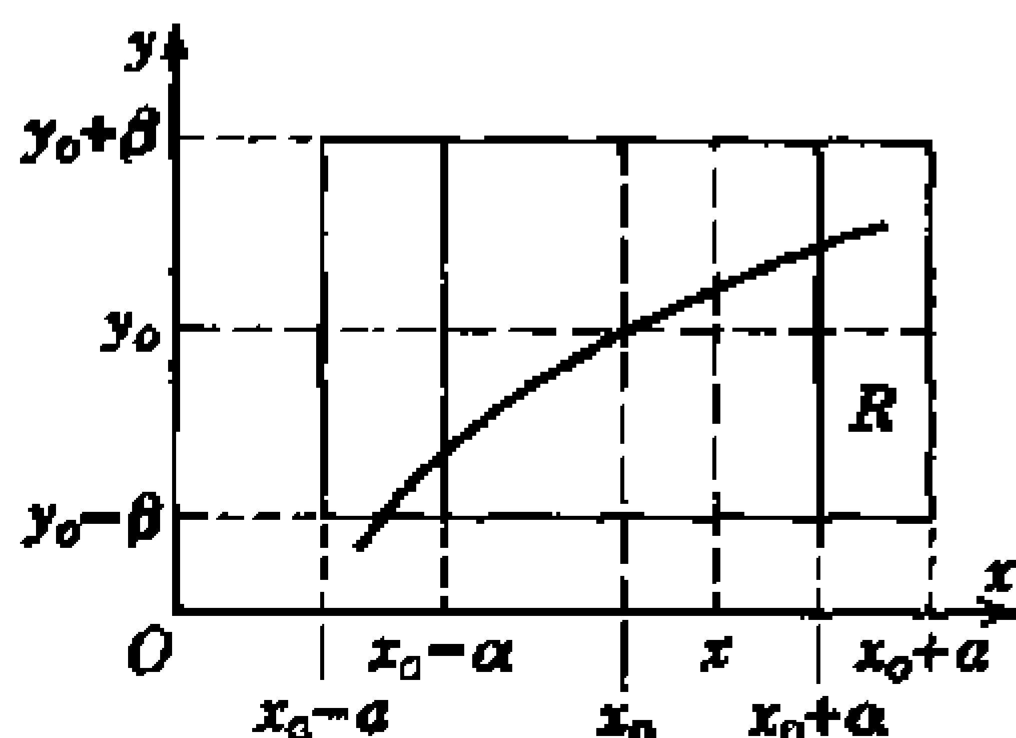


图 3.4

的. 这样, 对于所有 R 中的 (x,y) 存在正常数 m, M 使得

$$F_y(x,y) > m/2, \quad |F_x(x,y)| \le M. \tag{6}$$

对于由 $x_0 - a$ 到 $x_0 + a$ 之间的任一固定的 x, 表达式 $F(x,y)$ 在区间 $y_0 - \beta \le y \le y_0 + \beta$ 上是 y 的连续而单调递增的函数. 如果

$$F(x, y_0 + \beta) > 0, \quad F(x, y_0 - \beta) < 0, \tag{7}$$

则我们可以断定在由 $y_0 - \beta$ 到 $y_0 + \beta$ 之间存在一个唯一的值 y 满足 $F(x,y) = 0$. 于是对于给定 x, 方程 $F(x,y) = 0$ 有一唯一解 $y = f(x)$ 满足

$$y_0 - \beta < y < y_0 + \beta.$$

为了证明 (7), 我们看到, 按中间值定理

$$F(x, y_0) = F(x, y_0) - F(x_0, y_0) = F_x(\xi, y_0)(x - x_0),$$

这里 ξ 是 x_0 与 x 之间一个中间值. 因此, 如 α 表示 0 与 a 之间的一个数, 则我们就有

$$|F(x, y_0)| \le |F_x(\xi, y_0)|\ |x - x_0| \le M\alpha, \ |x - x_0| < \alpha.$$

相似地, 从 $F_y > m/2$ 推得

$$F(x, y_0 + \beta) = [F(x, y_0 + \beta) - F(x, y_0)] + F(x, y_0)$$

$$> \frac{1}{2}m\beta - M\alpha,$$

$$F(x, y_0 - \beta) = -[F(x, y_0) - F(x_0, y_0 - \beta)] + F(x, y_0)$$

$$< -\frac{1}{2}m\beta + M\alpha.$$

于是，只要我们取 α 相当小使得 $\alpha \le a$ 与 $\alpha < m\beta/2M$，那么对区间 $x_0 - \alpha \le x \le x_0 + \alpha$ 上的任何 x 不等式 (7) 都成立.

对于 $|x - x_0| \le \alpha$ 的任何 x，方程 $F(x,y) = 0$ 的解 $y = f(x)$ 在条件 $|y - y_0| \le \beta$ 与 $F_y(x,y) > m/2 > 0$ 之下，其存在性与唯一性便证明了. 而对 $x = x_0$，方程 $F(x,y) = 0$ 有相当于我们的初始点的解 $y = y_0$. 因为 y_0 必然落在 $y_0 - \beta$ 到 $y_0 + \beta$ 之间，因而我们有 $f(x_0) = y_0$. 而 $f(x)$ 的连续性与可微性现在可由多元函数的中值定理应用于 $F(x,y)$ 推出来 [见第 74 页 (33)]. 设 x 与 $x + h$ 系在 $x_0 - \alpha$ 到 $x_0 + \alpha$ 之间的两个值. 又设 $y = f(x)$ 与 $y + k = f(x + h)$ 是 f 的相应的值，这里 $y = f(x)$ 与 $y + k = f(x + h)$ 都位于 $y_0 - \beta$ 到 $y_0 + \beta$ 之间. 于是 $F(x,y) = 0$, $F(x + h, y + k) = 0$. 由此推出

$$0 = F(x + h, y + k) - F(x, y)$$

$$= F_x(x + \theta h, y + \theta k)h + F_y(x + \theta h, y + \theta k)k,$$

这里 θ 是 0 到 1 之间某一适当的中间值.[1]

应用 $F_y \ne 0$ 这一条件，我们可以用 F_y 来除并求得

$$\frac{k}{h} = -\frac{F_x(x + \theta h, y + \theta k)}{F_y(x + \theta h, y + \theta k)}. \tag{8}$$

由于矩形中的所有点都有 $|F_x| \le M$, $|F_y| > m/2$, 故可知上式右边有界 $2M/m$. 因此,

$$|k| \le \frac{2M}{m}|h|.$$

1) 注意中值定理在这里可以应用，因为连接矩形 $|x - x_0| \le \alpha$, $|y - y_0| \le \beta$ 内任意两点的线段整个落在矩形之内.

因此，当 $h \to 0$ 时 $k = f(x+h) - f(x) \to 0$，而这就表示 $y = f(x)$ 是连续函数. 由 (8) 推出，对固定的 x 并对 $y = f(x)$ 有

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = -\lim_{h\to 0} \frac{F_x(x+\theta h, y+\theta k)}{F_y(x+\theta h, y+\theta k)}$$
$$= -\frac{F_x(x,y)}{F_y(x,y)}.$$

这就建立了 f 的可微性，并且同时又给出微商公式 (4).

这个证明关键是 $F_y(x_0, y_0) \neq 0$ 这个假设，由此我们才能够推得 F_y 在 (x_0, y_0) 的一个足够小的邻域内不变号，并且 $F(x,y)$ 对固定的 x 是 y 的单调函数.

这个证明仅告诉我们函数 $y = f(x)$ 是存在的. 它是一个纯粹"存在性定理"的典型例子，其中计算解的实际可能性是不考虑的. 当然我们可以引用第一卷中讲到的任何数值方法 (见第一卷 524 页以下) 来逼近方程 $F(x,y) = 0$ 在固定的 x 处的解.

练 习 3.1 d

1. 给出一个函数 $f(x,y)$ 的例子，使得

a) 在 $x = x_0$, $y = y_0$ 的附近，$f(x,y) = 0$ 可以对 y 解出成为 x 的一个函数，

b) $f_y(x_0, y_0) = 0$.

2. 给出一个方程 $F(x,y) = 0$ 的例子，使在一点 (x_0, y_0) 的附近可以解出 y 为 x 的一个函数 $y = f(x)$，而 f 在 x_0 不是可微的.

3. 设 $\phi(x)$ 对 x 的所有实值都有定义. 试证明方程

$$F(x,y) = y^3 - y^2 + (1+x^2)y - \phi(x) = 0$$

对 x 的每一 x 的值都定义一个单值的 y.

e. 多于两个自变量的隐函数定理

隐函数定理可以扩充到多个自变数的函数如下：

设 $F(x,y,\cdots,z,u)$ 是自变数 $(x,y,\cdots,z,u)$ 的连续函数，并有连续的偏微商 $F_x,F_y,\cdots,F_z,F_u$. 设 $(x_0,y_0,\cdots,z_0,u_0)$ 为 F 的定义域的内点，在这儿，

$$F(x_0,y_0,\cdots,z_0,u_0)=0, \quad 及 \quad F_u(x_0,y_0,\cdots,z_0,u_0)\neq 0.$$

由是，我们可以对 u_0 划出一个区间 $u_0-\beta\leq u\leq u+\beta$ 和一个包含在其内部的点 $(x_0,y_0,\cdots,z_0)$ 的矩形区域 R，使得对 R 内每一个点 $(x,y,\cdots,z)$，方程 $F(x,y,\cdots,z,u)=0$ 被位于区间 $u_0-\beta\leq u<u_0+\beta$ 内的 u 的一个单一值所满足.[1] 对 u 的这个值，我们用 $u=f(x,y,\cdots,z)$ 表示，方程

$$F(x,y,\cdots,z,f(x,y,\cdots,z))=0$$

就成为在 R 内的恒等式;此外

$$u_0=f(x_0,y_0,\cdots,z_0),$$

$$u_0-\beta<f(x,y,\cdots,z)<u_0+\beta;$$

$$F_u(x,y,\cdots,z,f(x,y,\cdots,z))\neq 0.$$

函数 f 是自变数 $x,y,\cdots,z$ 的连续函数，并且具有连续的偏微商由下列方程给出:

$$F_x+F_uf_x=0,\ F_y+F_u\cdot f_y=0,\cdots,F_z+F_uf_z=0 \tag{9a}$$

证明完全与上节对于方程 $F(x,u)=0$ 的解所给出的证明遵循同一思路而没有新的困难.

有启发性的是，把微分公式 (9a) 结合成一个单一的方程

$$F_xdx+F_ydy+\cdots+F_zdz+F_udu=0. \tag{9b}$$

1) 值 β 与矩形域 R 并非唯一确定的. 如果我们取 β 为任何一个充分小的正数并且取 R(依赖于 β) 也充分小，定理的论断仍然是正确的.

用话来说，如果自变量 $x, y, \cdots, z, u$ 不是彼此独立而由条件 $F(x, y, \cdots, z, u) = 0$ 所约束，那么这些自变量的改变量的线性部分同样不是独立的，而由线性方程

$$dF = F_x dx + F_y dy + \cdots + F_z dz + F_u du = 0$$

连结在一起.

如果我们在 (9b) 内把 du 用表达式 $u_x dx + u_y dy + \cdots + u_z dz$ 代替，再让每个自变量的微分 $dx, dy, \cdots, dz$ 的系数等于 0，我们就得到微分公式 (9a).

顺便说一下，隐函数的概念使我们能够给出**代数函数**的一般定义. 我们称 $u = f(x, y, \cdots, z)$ 是自变量 $x, y, \cdots, z$ 的代数函数，如果 u 可以由方程 $F(x, y, \cdots, z, u) = 0$ 来隐式表示，其中 F 是变量 $x, y, \cdots, z, u$ 的一个多项式；简单地说，就是 u 满足一个**代数方程**. 一个不满足任何代数方程的函数称为**超越的**.

作为一个例子，我们把微分公式运用于球面的方程

$$F(x, y, u) = x^2 + y^2 + u^2 - 1 = 0.$$

求偏微商，我们得到

$$u_x = -\frac{x}{u}, \quad u_y = -\frac{y}{u};$$

再进一步微分得到

$$\begin{aligned}
u_{xx} &= -\frac{1}{u} + \frac{x}{u^2} u_x = -\frac{x^2 + u^2}{u^3}, \\
u_{xy} &= \frac{x}{u^2} u_y = -\frac{xy}{u^3}, \\
u_{yy} &= -\frac{1}{u} + \frac{y}{u^2} u_y = -\frac{y^2 + u^2}{u^3}.
\end{aligned}$$

练　习　3.1 e

1. 求证方程 $x + y + z = \sin xyz$ 可以在原点 $(0, 0, 0)$ 附近对 z

解出. 求出解的偏微商.

2. 对下列每一方程, 在所指出点附近考察是否都能唯一地对 z 解出而成为其余变量的函数:

(a) $\sin x+\cos y+\operatorname{tg} z=0\left(x=0, y=\frac{\pi}{2}, z=\pi\right)$;

(b) $x^2+2y^2+3z^2-\omega=0$

$(x=1, y=2, z=-1, \omega=8)$;

(c) $1+x+y=\cosh(x+z)+\sinh(y+z)$

$(x=y=z=0)$.

3. 求证 $x+y+z+xyz^3=0$ 在原点 $(0,0,0)$ 附近隐式地定义 z 为 x, y 的一个函数. 把 z 展开成为 x, y 的幂, 直到四阶为止.

3.2 用隐函数形式表出的曲线与曲面

a. 用隐函数形式表出的平面曲线

用形式为 $y=f(x)$ 的方程来描绘平面曲线时, 坐标间有主从区别, 并不对称. 曲线的 **切线** 与 **法线** 分别由方程

$$(\eta-y)-(\xi-x)f'(x)=0. \tag{10a}$$

与

$$(\eta-y)f'(x)+(\xi-x)=0 \tag{10b}$$

给出 (见第一卷第 386 页). 这里 ξ, η 是法线或切线上一任意点的 "流动坐标", 而 x, y 是曲线上的点的坐标. 曲线的 **曲率** 为

$$\kappa=\frac{f''}{(1+f'^2)^{3/2}}. \tag{10c}$$

(见第一卷第 39 页). 而对拐点则要求条件

$$f''(x)=0 \tag{10d}$$

成立. 我们现在将导出关于用 $F(x,y)=0$ 形式的方程隐式地表示的曲线的一系列相应的对称公式. 我们假定在所考虑的点处, F_x, F_y 不都是 0, 即

$$F_x^2+F_y^2\neq 0. \tag{11}$$

譬如假定 $F_y\neq 0$, 我们就可以在 (10a, b) 中对 $f'(x)$ 用 (4)(第 225 页) 的值来代替, 从而马上得到切线方程的另一个形式

$$(\xi-x)F_x+(\eta-y)F_y=0 \tag{12a}$$

与法线方程的相应形式

$$(\xi-x)F_y-(\eta-y)F_x=0. \tag{12b}$$

对 $F_y=0$, 而 $F_x\neq 0$, 我们则通过隐式方程 $F(x,y)=0$ 的另一种形式 $x=g(y)$ 得到相同的方程.

曲线在点 (x,y) 处的法线方向余弦, 即 (ξ,η) 平面上垂直于以 (12a) 为其方程的直线的方向余弦由

$$\cos\alpha=\frac{F_x}{\sqrt{F_x^2+F_y^2}},\quad \sin\alpha=\frac{F_y}{\sqrt{F_x^2+F_y^2}} \tag{12c}$$

给出 (见第 146 页 (20)). 类似地, 曲线的切线的方向余弦 —— 即直线 (12b) 的法线 —— 是

$$\cos\beta=\frac{-F_y}{\sqrt{F_x^2+F_y^2}},\quad \sin\beta=\frac{F_x}{\sqrt{F_x^2+F_y^2}}. \tag{12d}$$

曲线在一给定点的法线实际上有两个方向, 一个是由 (12c) 给出的方向余弦方向, 另一个与之相反. 由 (12c) 给出的法线与以 F_x, F_y 作为分量的向量,即 F 的梯度 (见第 222 页) 都是同向的. 由于梯度向量的方向就是 F 增加最快的那个方向,所以在曲线 $F(x,y)=0$ 的一点处的梯度总是指向区域 $F>0$ 的内部的. 对于公式 (12c) 所确定的法线方向也有同样的结论.

第 242 页公式 (5) 给出了由隐函数形式 $F(x,y)=0$ 给出的函数 $y=f(x)$ 的二阶微商 $y''=f''(x)$ 的表示式. 对于隐式 $F(x,y)=0$ 给出的曲线, 拐点出现的必要条件 $f''=0$ 可以写成

$$F_y^2F_{xx}-2F_xF_yF_{xy}+F_x^2F_{yy}=0. \tag{13}$$

在这公式中两个变量 x,y 没有那个是为主的. 现在它完全是对称的, 所以不再需要 $F_y\neq 0$ 的假定. 当然, 这个对称特征反映了拐点这个概念有完全与任何坐标系无关的几何意义.

若在曲线的曲率公式 (10c) 中用 $f''(x)$ 的公式 (5) 代入, 那么我们又得到一个 x 与 y 对称的表达式[1)]

$$\kappa=\frac{F_y^2F_{xx}-2F_xF_yF_{xy}+F_x^2F_{yy}}{\left(\sqrt{(F_x^2+F_y^2)}\right)^3}. \tag{14a}$$

引进**曲率半径**

$$\rho=\frac{1}{\kappa}, \tag{14b}$$

我们就得到**曲率中心**—— 即与 (x,y) 距离为 ρ 的内法线上的点 (见第一卷第 403 页)(ξ,η) 的坐标:

$$\xi=x-\frac{\rho F_x}{\sqrt{F_x^2+F_y^2}},\quad \eta=y-\frac{\rho F_y}{\sqrt{F_x^2+F_y^2}}. \tag{14c}$$

如果代替曲线 $F(x,y)=0$, 我们考虑曲线 $F(x,y)=c$, 其中 c 为常数, 则所有上面的讨论都一样成立. 我们只需用 $F(x,y)-c$ 代替函数 $F(x,y)$, 而这两者有着相同的导函数. 所以对这些曲线来说, 切线、法线以及还有一些其他对象的方程在形式上同上面完全一样.

我们让 c 取遍一个区间内的所有值, 便得到由 $F(x,y)-c=0$ 表示的曲线族, 它构成函数 $F(x,y)$ 的等高线族 (见第 16 页). 更一

1) 曲率的符号见第一卷第 400 页. 由 (14a) 定义的曲率是正的, 如果沿曲线的"外侧"F 在增加, 也就是说, 在接触点附近曲线的切线是位于 $F\geq 0$ 区域的.

般地说来，我们由形式为

$$F(x,y,c)=0$$

的方程得到一个**单参量的曲线族**. 当 c 取一给定值时，就得到一条等高线 Γ_c 的隐式表示. 对于位于 Γ_c 上的 —— 即满足方程

$$F(x,y,c)=0$$

的一点 (x,y), 以上推导出的公式都完全适用. 特别要提到的是，梯度向量 $(F_x(x,y,c),F_x(x,y,c))$ 在点 (x,y) 处与 Γ_c 垂直.

作为一个例子，我们来考虑椭圆

$$F(x,y)=\frac{x^2}{a^2}+\frac{y^2}{b^2}=1. \tag{15a}$$

由 (12a), 在点 (x,y) 的切线方程是

$$(\xi-x)\frac{x}{a^2}+(\eta-y)\frac{y}{b^2}=0.$$

因此，结合 (15a), 这可简化为

$$\frac{\xi x}{a^2}+\frac{\eta y}{b^2}=1.$$

由 (14a) 得到曲率

$$\kappa=\frac{a^4b^4}{(a^4y^2+b^4x^2)^{3/2}}. \tag{15b}$$

如果 $a>b$, 这在顶点 $y=0$, $x=\pm a$ 处有最大值 a/b^2, 而在另一对顶点 $x=0$, $y=\pm b$ 处则出现最小值 b/a^2.

如果两条曲线 $F(x,y)=0$ 与 $G(x,y)=0$ 在点 (x,y) 相交，**曲线的交角**是用交点处的两条切线 (或法线) 的交角 ω 来定义的. 如果回忆到法线方向可由梯度给出，并对这两个向量的夹角引用第 138 页公式 (7), 我们就得到公式

$$\cos\omega=\frac{F_xG_x+F_yG_y}{\sqrt{F_x^2+F_y^2}\sqrt{G_x^2+G_y^2}}. \tag{16}$$

这里 $\cos\omega$ 可按 F,G 增加的方向取定两条曲线的法线的交角 ω 来唯一地确定.

在 (16) 中令 $\omega=\dfrac{\pi}{2}$, 即在 (x,y) 点处曲线以直角相交, 就得到**正交性**的条件:

$$F_x\cdot G_x+F_y\cdot G_y=0. \tag{16a}$$

如果曲线**密接**, 即在交点处有一公共的切线与法线, 那么梯度向量 (F_x,F_y) 与 (G_x,G_y) 必相平行, 这就导致平行条件

$$F_xG_y-F_yG_x=0. \tag{16b}$$

作为一个例子, 我们考虑所有以原点作为焦点的抛物线族 ("共焦点" 抛物线)

$$F(x,y,c)=y^2-2c\left(x+\frac{c}{2}\right)=0 \tag{17a}$$

(见第 267 页图 3.9). 如果 $c_1>0$ 与 $c_2<0$ 则两支抛物线

$$F(x,y,c_1)=y^2-2c_1\left(x+\frac{c_1}{2}\right)=0$$

与

$$F(x,y,c_2)=y^2-2c_2\left(x+\frac{c_2}{2}\right)=0$$

正交于两个点; 因为在两个交点处有

$$x=-\frac{1}{2}(c_1+c_2),\quad y^2=-c_1c_2,$$

因之

$$F_x(x,y,c_1)F_x(x,y,c_2)+F_y(x,y,c_1)\cdot F_y(x,y,c_2)$$

$$=4(c_1c_2+y^2)=0.$$

由 (14a), 抛物线 (17a) 的曲率是

$$\kappa=\frac{c^2}{(c^2+y^2)^{3/2}}.$$

在顶点

$$(x,y)=\left(-\frac{c}{2},0\right)$$

处就简化为

$$\kappa=\frac{1}{|c|}.$$

在顶点处曲率中心或密切圆的中心，由 (14c) 而有坐标

$$\xi=-\frac{c}{2}+|c|\operatorname{sgn} c=\frac{c}{2},\quad \eta=0;$$

因此焦点 (0,0) 就是顶点与曲率中心的中点.

练 习 3.2 a

1. 求下列隐示地给出的曲线的切线与法线方程:

(a) $x^2+2y^2-xy=0$,

(b) $e^x\sin y+e^y\cos x=1$,

(c) $\cosh(x+1)-\sin y=0$,

(d) $x^2+y^2=y+\sin x$,

(e) $x^3+y^4=\cosh y$,

(f) $x^y+y^x=1$.

2. 计算曲线

$$\sin x+\sin y=1$$

在原点处的曲率.

3. 求极坐标方程 $f(r,\theta)=\theta$ 所给出的曲线的曲率.

4. 证明曲线

$$(x+y-a)^3+27axy=0$$

与直线 $x+y=a$ 的交点是曲线的拐点.

5. 确定常数 a 与 b 使得圆锥曲线

$$4x^2+4xy+y^2-10x-10y+11=0,$$

$$(y+bx-1-b)^2-a(by-x+1-b)=0$$

在点 (1,1) 处互相正交并且在该点有相同的曲率.

6. 设 K' 和 K'' 是两个圆, 有两个交点 A 与 B, 如果另一圆 K 与 K' 和 K'' 均正交, 那么他也必与每一个经过交点 A 与 B 的圆相正交.

b. 曲线的奇点

在前一节的许多公式中, 都在分母中出现表达式 $F_x^2+F_y^2$. 于是, 我们可以推测在 $F_x^2+F_y^2=0$ 的那些点处可能会出现某种奇异现象. 在曲线 $F(x,y)=0$ 的这样一个点处 $F_x=0$ 与 $F_y=0$, 因而作为切线的斜率的表达式 $y'=-\dfrac{F_x}{F_y}$ 失掉了意义.

如果在一个点 P 的邻域内, 一个变量 x 或是 y 可以表成另一个变量的连续可微函数, 则称该点 P 为曲线的一个正规点. 在这种情况下, 曲线在点 P 有切线并在点 P 的附近与切线相差很小. 曲线上正规点之外的点都称为奇点.

从隐函数定理我们知道, 若 $F(x,y)$ 有连续的一阶偏微商, 则在曲线上每一个适合条件 $F_x^2+F_y^2\neq 0$ 的点都是正规点; 因为在 P 点处若 $F_y\neq 0$, 我们可以由方程 $F(x,y)=0$ 解出一个连续可微的解 $y=f(x)$, 而若 $F_x\neq 0$ 则可以按 x 解出方程.

有各种类型的奇点. 其中重要的有所谓重点, 即有两个或更多的曲线分支通过的点. 例如双纽线 (第一卷第 113 页)

$$(x^2+y^2)^2-2a^2(x^2-y^2)=0$$

即以原点为一个重点. 很清楚, 在重点的邻域内不能把曲线的方程唯一地表成 $y=f(x)$ 或 $x=g(y)$ 的形式.

作为非重点的奇点的一个例子，可举出半立方曲线

$$F(x,y)=y^3-x^2=0$$

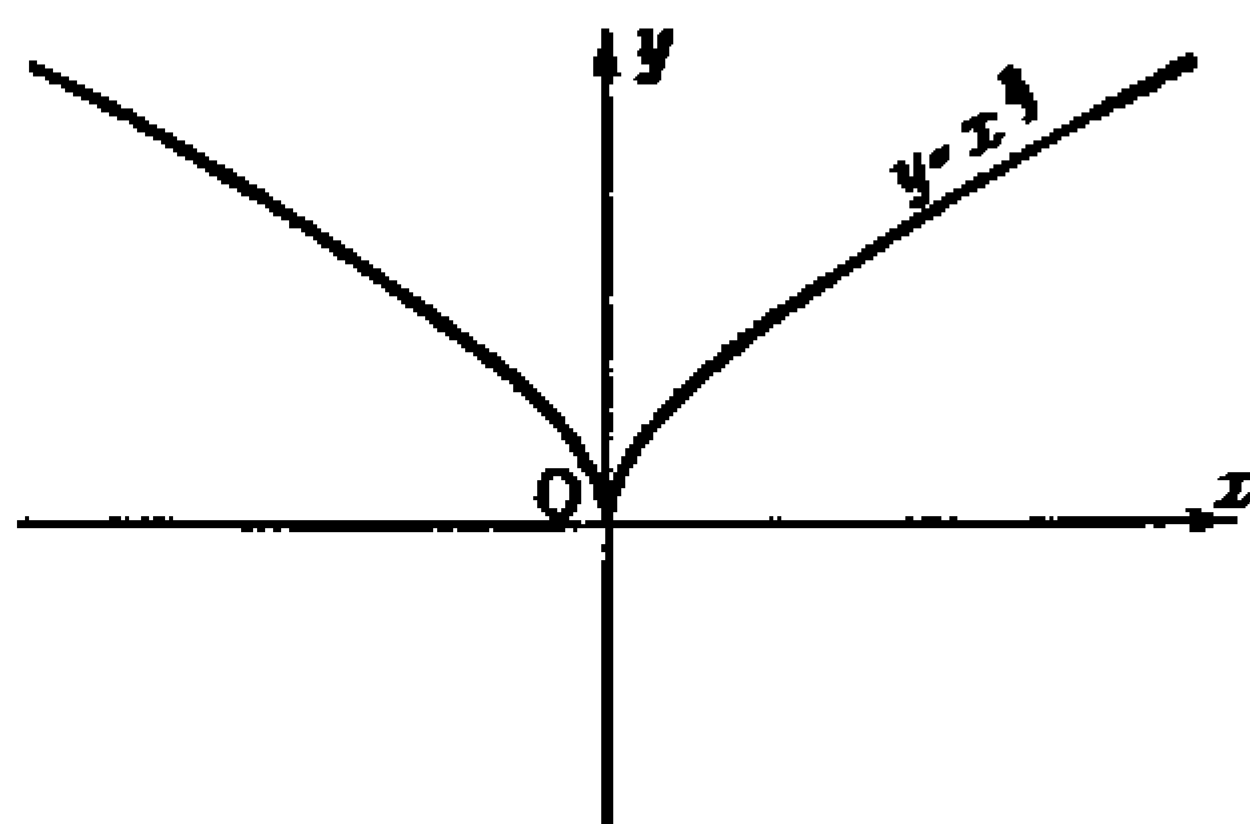


图 3.5　曲线 $y^3-x^2=0$

(见图 3.5). 这在原点处有 $F_x=F_y=0$. 解出 y, 则可以把曲线方程表成

$$y=f(x)=\sqrt[3]{x^2},$$

其中 f 是连续的，但在原点不可微．这曲线在原点有一个尖点．

有时 F_x 和 F_y 在一点处皆为 0, 曲线在该点仍可能是正规的．这可以

$$F(x,y)=y^3-x^4=0$$

为例，这在原点处有 $F_x=F_y=0$. 对 y 求解即得

$$y=f(x)=\sqrt[3]{x^4},$$

其中 $f(x)$ 对所有 x 都是连续可微的; 因此原点确实是一个正规点．由于 F 是 x 的偶函数，曲线对于 y 轴对称、曲线是凹的，并与 x 轴在原点相密接，就好像抛物线 $y=x^2$ 那样．但是原点也是曲线的某种特殊点，因为 f'' 在这里变成无穷大，即曲线在此取**无穷大曲率**．

由代表直线 $y=x$ 的这个方程

$$F(x,y)=(y-x)^2=0$$

这一简单例子可以看到，曲线 $F(x,y)=0$ 就是在 $F_x^2+F_y^2=0$ 的点处也无任何奇异的性态.

我们将在附录 A.3 中更系统地研究曲线的奇点.

练 习 3.2 b

1. 对下列曲线在原点处讨论其奇异性:

(a) $F(x,y)=ax^3+by^3-cxy=0$,

(b) $F(x,y)=(y^2-2x^2)^2-x^5=0$,

(c) $F(x,y)=(1+e^{1/x})y-x=0$,

(d) $F(x,y)=y^2(2a-x)-x^3=0$,

(e) $F(x,y)=(y-2x)^2-x^5=0$.

2. 曲线 $x^3+y^3-3axy=0$ 在原点有一双重点，问其切线是什么?

3. 画出曲线 $(y-x^2)^2-x^5=0$ 的图形并证明它在原点处有一尖点. 与曲线 $y^3-x^2=0$ 的尖点比较，这个尖点有何特殊之处?

4. 试证明曲线族

$$(x\cos\alpha-y\sin\alpha-b)^3=c(x\sin\alpha+y\cos\alpha)^2$$

(其中 α 是参数，b,c 是常数) 的每一条都有一个尖点，并且全部尖点都位于同一圆周上.

5. 设 (x,y) 是曲线 $F(x,y)=0$ 的一个双重点. 假定在 (x,y) 处并非所有二级微商全为 0, 试计算出 (x,y) 处的两条切线的交角 ϕ. 求出

(a) 双纽线,

(b) 笛卡儿叶形线 (见第 243 页)

的双重点处的两切线的交角.

6. 试求出曲线

$$y(ax+by)=cx^3+ex^2y+fxy^2+gy^3$$

在原点处两个分支的每一个的曲率.

c. 曲面的隐函数表示法

直到现在，我们常用函数 $z=f(x,y)$ 形式来表示 x,y,z 空间的曲面. 对于空间中某一给定的曲面的这种显式表示中所给予 z 的这种特殊地位，人们常感到不方便. 比较自然也比较一般的是用 $F(x,y,z)=0$ 或 $F(x,y,z)=c$ 这种隐式来表示空间中的曲面. 例如以原点为中心的球面，用对称形式方程

$$x^2+y^2+z^2-r^2=0$$

来表示比用方程

$$z=\pm\sqrt{r^2-x^2-y^2}$$

更好. 曲面的显式表示也可看作一种特殊的隐式表示，即

$$F(x,y,z)=z-f(x,y)=0.$$

为要导出曲面 $F(x,y,z)=0$ 在点 P 处的切平面方程，我们作一个假定，即在该点处

$$F_x^2+F_y^2+F_z^2\neq 0, \tag{18}$$

也就是至少有一个偏微商不是 0[1]. 比如说，$F_z\neq 0$, 我们就可以找到曲面在 P 点附近的显式方程 $z=f(x,y)$. 在 P 点处的切平面的方程便是

$$\zeta-z=(\xi-x)f_x+(\eta-y)f_y, \tag{19a}$$

以 (ξ,η,ζ) 为流动坐标 (见第 52 页).

按第 248 页公式 (9a), 用 $f_x=-F_x/F_z$, $f_y=-F_y/F_z$ 来代替 f 的偏微商，我们得到切平面方程，形为

$$(\xi-x)F_x+(\eta-y)F_y+(\zeta-z)F_z=0. \tag{19b}$$

---

1) 同曲线一样，梯度 F 的消失为 0 常相应于曲面的奇异性质，但我们不准备讨论这种奇点的性质.

切平面 (19b) 的法线与梯度向量 (F_x, F_y, F_z) 有相同的方向 (见第 222 页). 因此, 法线的方向余弦的表达式为

$$\cos\alpha = \frac{F_x}{\sqrt{F_x^2 + F_y^2 + F_z^2}},$$

$$\cos\beta = \frac{F_y}{\sqrt{F_x^2 + F_y^2 + F_z^2}},$$

$$\cos\gamma = \frac{F_z}{\sqrt{F_x^2 + F_y^2 + F_z^2}}.$$

更确切地说, 我们在这里取定了指向 F **增加**的那个方向作为平面的法线方向 (见第 223 页).

如果两曲面 $F(x,y,z) = 0$ 与 $G(x,y,z) = 0$ 在一点相交, **曲面的交角** ω 是由切平面的交角或是法线的交角来定义的. 即

$$\cos\omega = \frac{F_x G_x + F_y G_y + F_z G_z}{\sqrt{F_x^2 + F_y^2 + F_z^2} \cdot \sqrt{G_x^2 + G_y^2 + G_z^2}}. \tag{20a}$$

取其特例 $\left(\omega = \frac{\pi}{2}\right)$, 便得到正交条件

$$F_x G_x + F_y G_y + F_z G_z = 0. \tag{20b}$$

代替一个由 $F(x,y,z) = 0$ 方程给出的曲面, 我们可考虑由

$$F(x,y,z) = c$$

(此处 c 是一常数) 给出的更为广泛的曲面族. 常数 c 的不同值给出函数 F 的不同的等值面 (见 224 页). 在任何一点 (x,y,z) 处梯度向量 (F_x, F_y, F_z) 都是垂直于经过那点的等值面的. 同样地, 方程 (19b) 给出等值面的切平面.

作为一个例子, 我们考虑**球面**

$$x^2 + y^2 + z^2 = r^2.$$

由 (19b)，在点 (x,y,z) 处的切平面方程是

$$(\xi-x)2x+(\eta-y)2y+(\zeta-z)2z=0,$$

也就是

$$\xi x+\eta y+\zeta z=r^2.$$

法线的方向余弦是与 (x,y,z) 成比例的，也就是说，法线与由原点到点 (x,y,z) 的矢径是重合的.

对于以坐标轴为主轴的更一般的**椭球面**

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1,$$

切平面的方程是

$$\frac{\xi x}{a^2}+\frac{\eta y}{b^2}+\frac{\zeta z}{c^2}=1.$$

练 习 3.2 c

1. 找出下列曲面在指定点的切平面：

(a) 曲面

$$x^3+2xy^2-7z^3+3y+1=0$$

在点 (1,1,1) 处；

(b) 曲面

$$(x^2+y^2)^2+x^2-y^2+7xy+3x+z^4-z=14$$

在点 (1,1,1) 处；

(c) 曲面

$$\sin^2 x+\cos(y+z)=\frac{3}{4}$$

在点 $\left(\frac{\pi}{6},\frac{\pi}{3},0\right)$ 处；

(d) 曲面

$$1 + x\cos\pi z + y\sin\pi z - z^2 = 0$$

在点 (0,0,1) 处;

(e) 曲面

$$\cos x + \cos y + 2\sin z = 0$$

在点 $\left(0,0,-\frac{\pi}{2}\right)$ 处;

(f) 曲面

$$x^2 + y^2 = z^2 + \sin z$$

在点 (0,0,0) 处.

2. 证明曲面族

$$\frac{xy}{z} = u, \sqrt{x^2+z^2} + \sqrt{y^2+z^2} = v,$$

$$\sqrt{x^2+z^2} - \sqrt{y^2+z^2} = w$$

过同一点的三个曲面是互相正交的.

3. 设点 A 与 B 都是以同一速度作匀速运动, A 自原点沿 z 轴运动, B 从点 $(a,0,0)$ 沿平行于 y 轴的方向运动. 求由联线 AB 所生成的曲面.

4. 试证明曲面 $x^2+y^2-z^2=1$ 在任何一点的切平面都与曲面交于两条直线.

5. 如果 $F(x,y,z)=1$ 是曲面方程, F 为 n 次齐次函数, 那么点 (x,y,z) 的切平面由以下方程给出:

$$\xi F_x + \eta F_y + \zeta F_z = h.$$

6. 设 z 被方程

$$x^3 + y^3 + z^3 - 3xyz = 0$$

定义为 x,y 的函数, 试把 z_x, z_y 表为 x,y,z 的函数.

7. 求下列各对曲面在指定点的交角:

(a) $2x^4+3y^3-4z^2=-4$, $1+x^2+y^2=z^2$, 在点 (0, 0,1) 处;

(b) $x^y+y^2=2$, $\cosh(x+y-2)+\sinh(x+z-1)$, 在点 (1, 1, 0) 处;

(c) $x^2+y^2=e^z$, $x^2+z^2=e^y$, 在点 (1,0,0) 处;

(d) $1+\sinh(x/\sqrt{z})=\cosh(y/\sqrt{z})$, $x^2+y^2=z^2-1$, 在 (0,0,1) 点;

(e) $\cos\pi(x^2+y)+\sin(\pi(x^2+z))=1$, $x^3+y^3=z^3$, 在点 (0,0,0) 处.

3.3 函数组、变换与映射

a. 一般说明

由隐函数已得到的结果使得我们能够考虑函数 组, 即同时一起讨论几个函数. 这一节我们仅仅考虑特别重要的一种情形, 即函数的个数与自变量的个数恰好相等. 我们从研究两个自变量的这种函数组的意义开始. 假定两个函数

$$\xi=\phi(x,y) \quad 与 \quad \eta=\psi(x,y) \tag{21a}$$

都是在 xy 平面上的一个集合 R(函数的定义域) 上连续可微, 我们可以按两种途径来解释这个函数组. 第一种 (活动的) 的解释是作为 映射 或 变换.(第二种解释是看作坐标的变换, 将在第 268 页进行讨论.) 对于 xy 平面上以 (x,y) 为坐标的任意一点 P 在 $\xi\eta$ 平面上总是对应于一个以 (ξ,η) 为坐标的像点 Π.

一个简单的例子是所谓 仿射 映射或变换

$$\xi=ax+by, \quad \eta=cx+dy,$$

这里 a,b,c,d 都是常数 (见第 160 页).

通常把 (x,y) 与 (ξ,η) 解释为同一个平面上的点. 这时我们称之为 xy 平面到自身的映射或变换.

与映射相联系的一个基本问题是映射的逆问题, 即按方程组 $\xi=\phi(x,y)$ 与 $\eta=\psi(x,y)$ 如何把 x,y 表示为 ξ,η 的函数以及如何决定这些反函数的性质.

如果当 (x,y) 跑遍映射的定义域 R 时, 其像点跑遍 $\xi\eta$ 平面的一个集合 B, 我们就称 B 为 R 的像集或映射的值域. 如果 R 上的两个不同点总对应于 B 上两个不同的点, 则 B 上的每一个点 (ξ,η) 仅有 R 上一个唯一的点 (x,y) 把它作为像点. (点 (x,y) 叫做点 (ξ,η) 的逆象, 与像点相对立.) 这就是说, 我们可以把映射唯一地反转过来, 由在 B 上定义的函数组

$$x=g(\xi,\eta) \qquad y=h(\xi,\eta) \tag{21b}$$

来确定 x,y 作为 ξ,η 的函数. 这时我们说映射 (21a) 有唯一的逆或说它是一个 1-1 映射. 我们称变换 (21b) 为原来的逆映射或是逆变换.

如果在这个映射中, 点 $P(x,y)$ 描绘一条 R 内的曲线, 则像点 (ξ,η) 通常将相应地描绘出一条集合 B 内的曲线, 称为前一曲线的像曲线. 例如, 平行于 y 轴的直线 $x=c$ 在 $\xi\eta$ 平面上相应于下列以 y 作参量的这个参量方程组

$$\xi=\phi(c,y), \qquad \eta=\psi(c,y) \tag{22a}$$

所表示的曲线. 再则直线 $y=k$ 对应于曲线

$$\xi=\phi(x,k), \qquad \eta=\psi(x,k) \tag{22b}$$

如果对于 c 和 k 我们取定两串等距离的值 $c_1,c_2,c_3,\cdots$ 与 $k_1,k_2,k_3,\cdots$ 则由直线 $x=$ 常数与 $y=$ 常数组组成的矩形 "坐标网"(如同通常坐标纸上的直线网) 就对应于 $\xi\eta$ 平面上相应的曲线所组成的曲线坐

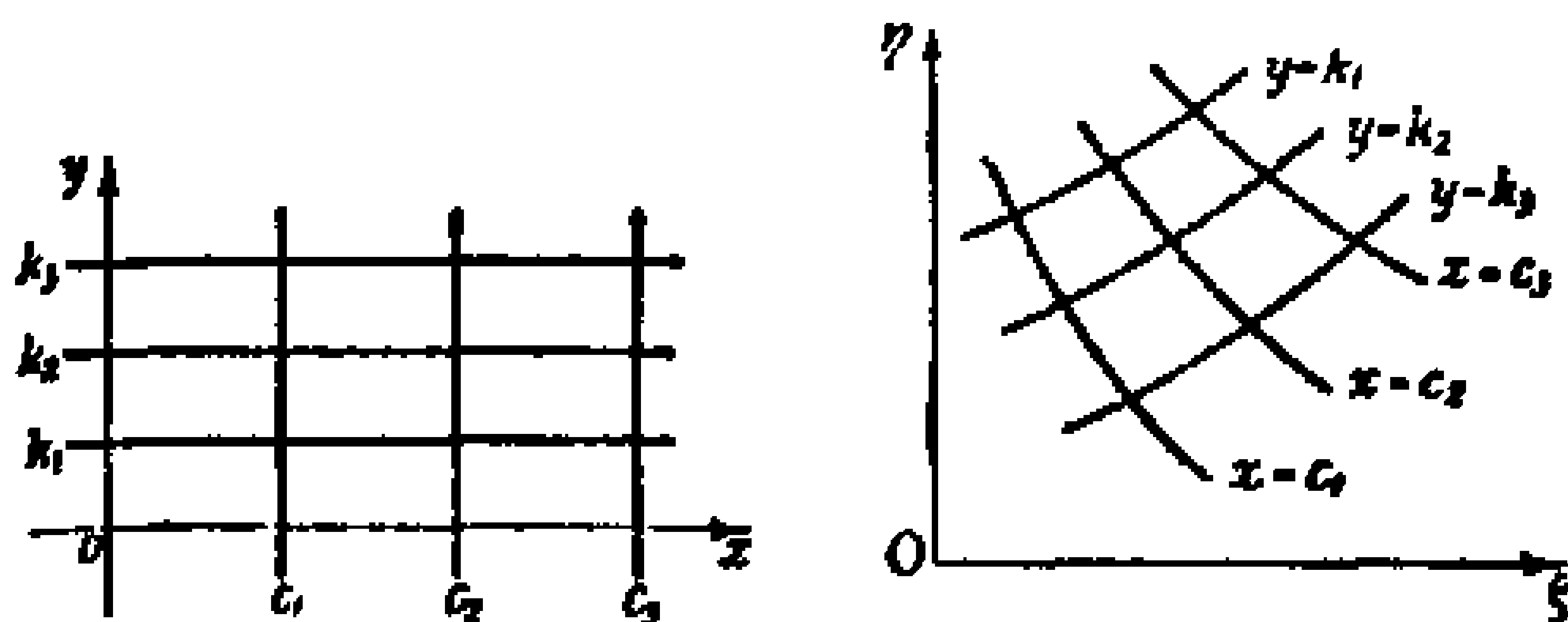


图 3.6 和图 3.7　在 xy 平面与 $\xi\eta$ 平面上 $x=$ 常数与 $y=$ 常数的曲线网

标网 (图 3.6 和图 3.7). 这两族曲线都可写成隐函数形式. 如果我们用方程 (21b) 表示逆变换; 那么曲线方程就分别是

$$g(\xi,\eta)=c, \qquad 与\ h(\xi,\eta)=k. \tag{22c}$$

在很多情况下, 曲线网给映射 (21a) 提供一个很有用的 **几何图象**, 比把该方程组解释为在 (x,y,ξ,η) 四维空间中的一个二维曲面要清晰得多.

同样地 $\xi\eta$ 平面上的两族直线 $\xi=r$ 与 $\eta=k$ 相应于 xy 平面上的两族曲线

$$\phi(x,y)=r \quad 与\ \psi(x,y)=k.$$

作为一个例子, 我们考虑 **反演**(又称 **逆半径映射** 或 **关于单位圆的反射**). 这个变换由下列方程组给出

$$\xi=\frac{x}{x^2+y^2}, \quad \eta=\frac{y}{x^2+y^2}. \tag{23a}$$

这里每一点 $P=(x,y)$ 都对应到一点 $\Pi=(\xi,\eta)$, 位于同一射线 OP 上, 并满足方程

$$\xi^2+\eta^2=\frac{1}{x^2+y^2} \quad 或 \quad OP=\frac{1}{O\Pi}; \tag{23b}$$

从而位移向量 $\overrightarrow{OP}$ 的长度是位移向量 $\overrightarrow{O\Pi}$ 的长度的倒数. 单位圆 $x^2+y^2=1$ 内部的点被映到单位圆外部的点, 反之也一样. 从 (23b)

我们求得其**逆变换**为

$$x = \frac{\xi}{\xi^2+\eta^2}, \qquad y = \frac{\eta}{\xi^2+\eta^2},$$

这也是一个**反演**; 因而反演变换下每一点的像点的逆与其像点是重合的.

作为映射 (23a) 的定义域 R, 我们可取除去原点外的全部平面, 而值域 B 则取除去原点外的整个 $\xi\eta$ 平面. 在 $\xi\eta$ 平面上的直线 $\xi=\gamma$ 与 $\eta=\kappa$ 分别对应于 xy 平面上的圆

$$x^2+y^2-\frac{1}{\gamma}x=0 \qquad \text{与} \qquad x^2+y^2-\frac{y}{\kappa}=0.$$

同样, xy 平面上的直线坐标网对应于两个圆周族, 分别在原点切于 ξ 轴和 η 轴.

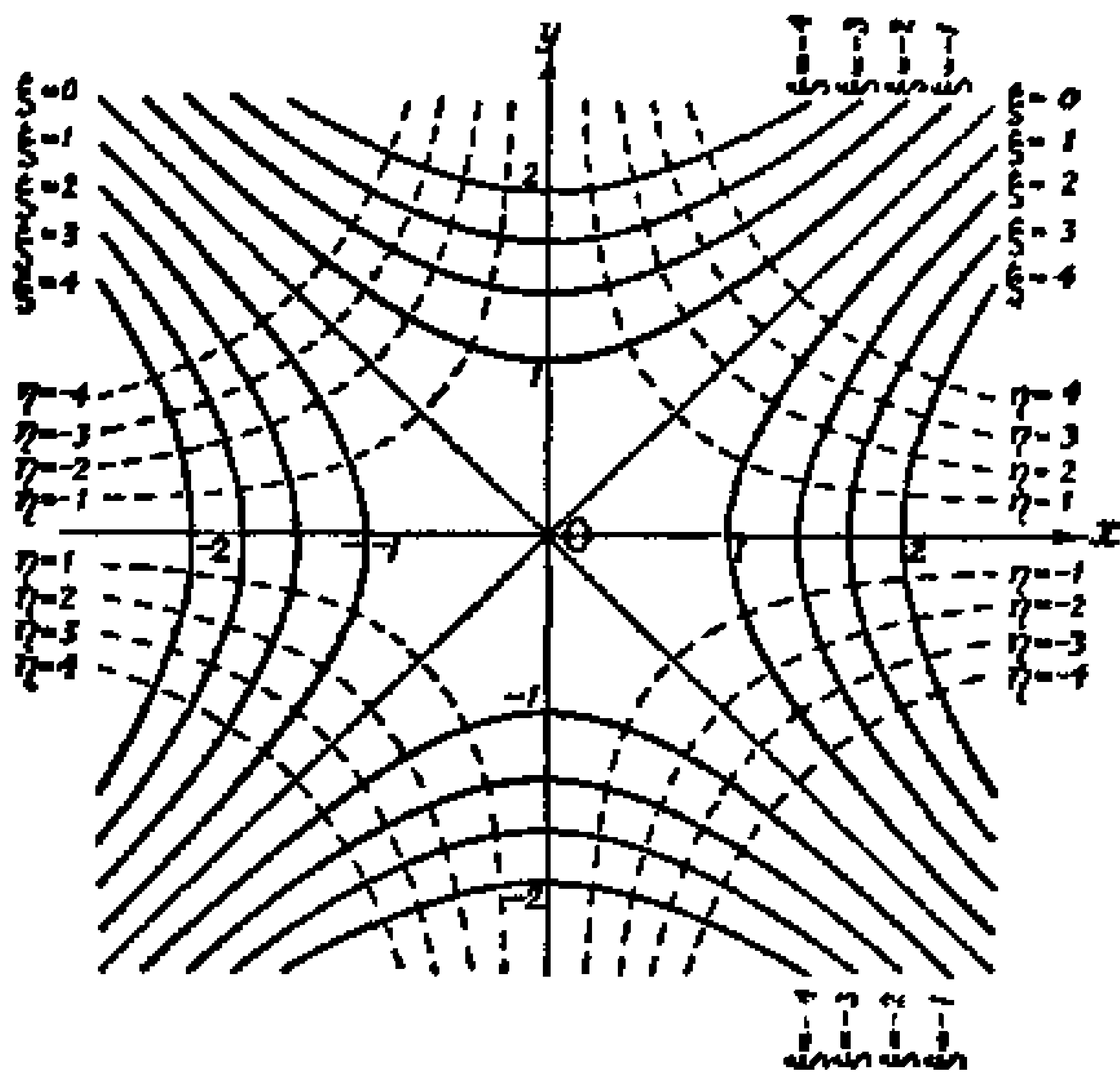


图 3.8

作为另一个例子, 我们考虑映射

$$\xi = x^2-y^2, \quad \eta = 2xy.$$

直线 $\xi=$ 常数映射到 xy 平面为等轴双曲线 $x^2-y^2=$ 常数，其渐近线为直线 $x=y$ 与 $x=-y$. 直线 $\eta=$ 常数也对应到一族等轴双曲线，以坐标轴为其渐近线. 每一族双曲线都与另一族双曲线交于直角 (图 3.8). 在 xy 平面上与坐标轴平行的直线对应到 $\xi\eta$ 平面上的两族抛物线：抛物线族 $\eta^2=4c^2(c^2-\xi)$ 对应到直线 $x=c$, 而抛物线族 $\eta^2=4\kappa^2(\kappa^2+\xi)$ 对应到直线 $y=\kappa$. 所有这些抛物线都以原点为焦点，而以 ξ 轴为其轴；它们形成一族共焦点与共轴的抛物线 (图 3.9)

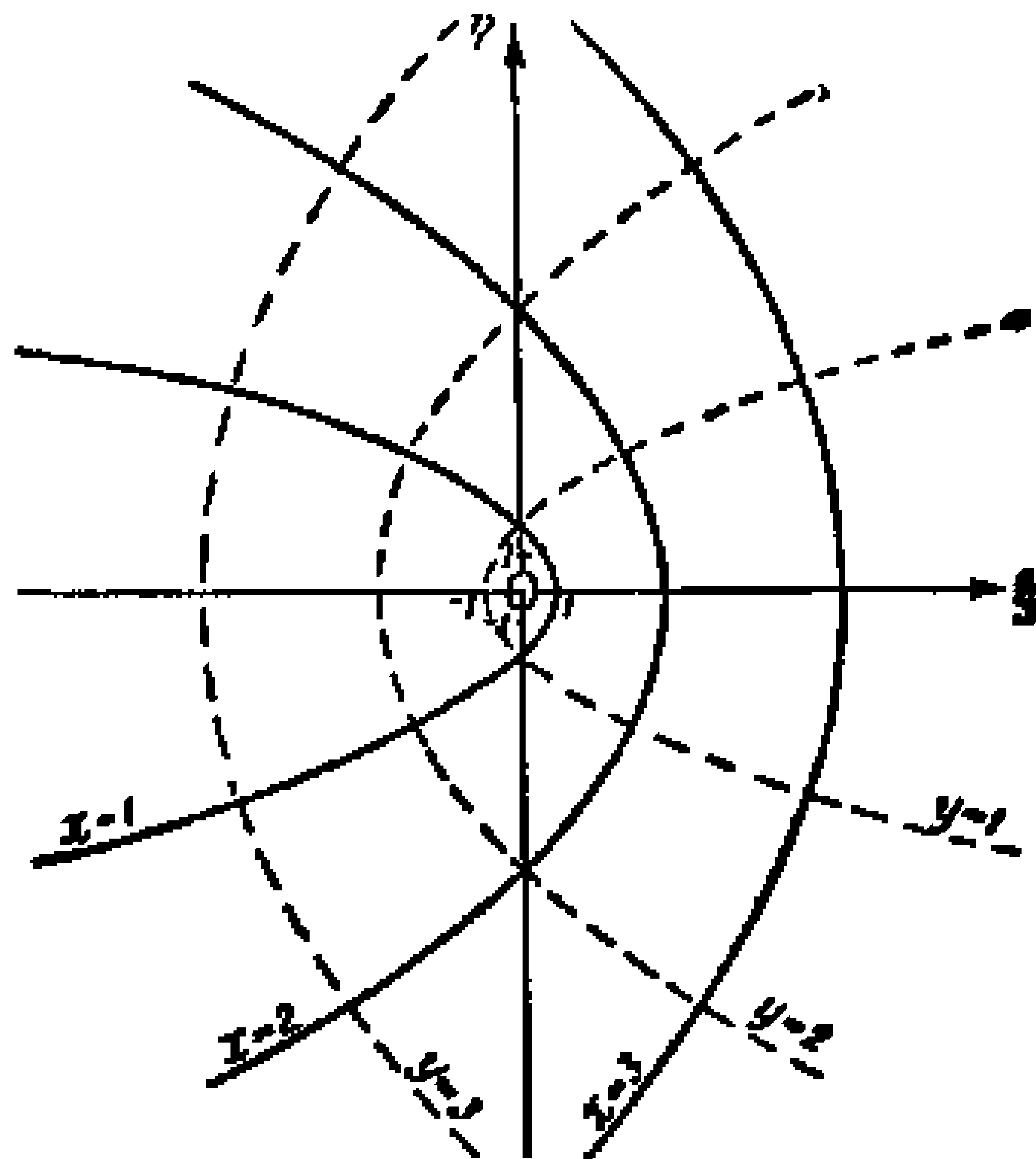


图 3.9

一一变换有一个重要的解释并应用于譬如流体这样的**连续介质的运动与变形**的描述. 如果我们考虑这样的介质在给定时间分布在区域 R 上，然后由一个运动使它变形，那原来分布在 R 上的介质将覆盖一个与 R 不同的区域 B. 在运动的开始介质的每一质点可以由 R 内的坐标 (x,y) 来区别，而在运动的结束用 B 内的坐标 (ξ,η) 来区别. 把 (x,y) 变到 (ξ,η) 的变换的 1 对 1 的特性不过是不同的质点还变为不同的质点这个显然的物理事实的数学表示而已.

练 习 3.3 a

1. 找出下列变换下直线 $x=$ 常数, $y=$ 常数的象曲线:

(a) $\xi = e^x \cos y$, $\eta = e^x \sin y$;

(b) $\xi = (x-y)/2$, $\eta = \sqrt{xy}$;

(c) $\xi = \sqrt{x/y}$, $\eta = \cos(x+y)$;

(d) $\xi = x + y^2$, $\eta = y + x^2 - 1$;

(e) $\xi = x^y$, $\eta = y^x$;

(f) $\xi = \sinh x$, $\eta = \cosh y$;

(g) $\xi = \sin(x+y)$, $\eta = \cos(x-y)$;

(h) $\xi = e^{\cos x}$, $\eta = e^{\sin y}$.

2. 找出在变换 $\xi = e^x, \eta = e^y$ 下由曲线 $\sinh^2 x + \cosh^2 y = 1$ 所包围的区域的像域.

3. 找出在变换 $\xi = \sqrt{x+y}, \eta = \sqrt{y-x}$ 下矩形域 $1 \le x \le 3$, $4 \le y \le 16$ 的像域.

4. 变换 $\xi = x - xy, \eta = 2xy$ 是否是一一的?

b. 曲线坐标

与函数组 $\xi = \phi(x,y), \eta = \psi(x,y)$ 的第一种解释 (看做映射) 紧密相联系的第二种解释是看成平面上的 **坐标变换**. 如果函数 ϕ, ψ 不是线性的, 则变换不再是仿射的, 而是 **一般的曲线坐标变换**.

我们再度明确假定, 当 (x,y) 跑遍 xy 平面上的区域 R 时, 相应点 (ξ,η) 跑遍 $\xi\eta$ 平面上的区域 B, 并且对 B 的每一点, 都能在 R 中唯一地确定一个相应点 (x,y); 换句话说, 变换是一一的. 它的反变换我们仍写成 $x = g(\xi,\eta)$, $y = h(\xi,\eta)$.

现在我们说区域 R 中一**点** P **的坐标**，我们的意思是指任何一个数对，它对于给定的一个坐标架能够唯一地确定点 P 在 R 中的位置. 直角坐标形成最简单的坐标系，它扩展到整个平面. 另一个熟知的坐标系是 xy 平面上的极坐标系，它由下列方程给出:

$$\xi = \gamma = \sqrt{x^2 + y^2}, \quad \eta = \theta = \operatorname{arctg} \frac{y}{x} \quad (0 \leq \theta < 2\pi).$$

当我们如上给定了一个函数组 $\xi = \phi(x, y)$, $\eta = \psi(x, y)$ 时，一般我们能够对每个点 $P(x, y)$ 指出相应值 (ξ, η) 作为它的新坐标，而对区域 B 上每一对值 (ξ, η) 唯一地确定数对 (x, y)，因而唯一地确定了点 P 在 R 上的位置. "坐标线"$\xi=$ 常数和 $\eta=$ 常数在 xy 平面上就代表两族曲线，分别由 $\phi(x, y) =$ 常数与 $\psi(x, y) =$ 常数两方程隐式地表示. 这些坐标曲线盖满区域 R 而成为一组坐标网 (通常为曲线) 的. 由于这个缘故，坐标 (ξ, η) 也称为 R 上的**曲线坐标**.

我们将再一次指出，方程组的这两种解释是如何紧密地相联系着的. 由 xy 平面上平行于坐标轴的直线映射到 $\xi\eta$ 平面上来的相应曲线，可以直接看做 $\xi\eta$ 平面上的曲线坐标系 $x = g(\xi, \eta)$, $y = h(\xi, \eta)$ 下的坐标曲线；反之，xy 平面上的曲线坐标系 $\xi = \phi(x, y)$, $\eta = \psi(x, y)$ 的坐标曲线，就是 $\xi\eta$ 平面上的与坐标轴平行的直线网在这映射中的像曲线. 即使在把 (ξ, η) 看做 xy 平面上的曲线坐标的解释中，如果我们要使情况理解得清楚的话，我们也必须考虑一个 $\xi\eta$ 平面与其上一个区域 B，使得坐标为 (ξ, η) 的点可以在其中变化. 区别主要在于观察的角度 [1). 如果我们主要兴趣在于 xy 平面上的区域 R，我们就把 (ξ, η) 看做用来确定区域 R 上的点的位置的一个新办法，而 $\xi\eta$ 平面上区域 B 就仅是次要的了；而如果我们对 xy 平面与 $\xi\eta$ 平面的两个区域 R 与 B 都同样有兴趣，那么我们最好是把方程组看做刻画两个区域之间的一个对应，也就是一个区域

---

1) 然而有一个真正的区别在于，这些方程永远定义了一个映射，不管有多少个点 (x, y) 对应到一个点 (ξ, η)；而这些方程却仅当对应是一一对应的时候，才定义一个坐标变换.

到另一个区域的一个映射. 不过, 往往最好是把两种解释 —— 映射与坐标变换都同时记在心上.

作为一个例子, 我们引进极坐标 (r,θ), 并把 r 与 θ 看做 $r\theta$ 平面上的直角坐标, 而圆周族 $r=$ 常数与直线族 $\theta=$ 常数则映射到 $r\theta$ 平面上的与坐标轴平行的直线. 如果 xy 平面上区域 R 是圆 $x^2+y^2\leq 1$, 那么 $r\theta$ 平面上的点 (r,θ) 将跑遍整个矩形 $0\leq r\leq 1$, $0\leq\theta\leq 2\pi$, 在这里 $\theta=0$ 和 $\theta=2\pi$ 这两条边上的点都对应到区域 R 的同一个点, 而整个边 $r=0$. 则是原点 $(x,y)=(0,0)$ 的象.

曲线坐标系的另一例子是**抛物线坐标系**. 我们从考虑 xy 平面上的共焦点的抛物线 (见第 267 页与图 3.9)

$$y^2=2c\left(x+\frac{c}{2}\right)$$

来得到它. 这族抛物线都以原点为焦点, 以 x 轴为轴. 通过平面上除原点外的每一点有抛物线族中的两条抛物线, 一条相应于正参数值 $c=\xi$, 而另一相应于负参数值 $c=\eta$. 我们用相应于点的这两个值 x,y 来解出方程

$$y^2=2c\left(x+\frac{c}{2}\right),$$

得到 c 的两个值:

$$\xi=-x+\sqrt{x^2+y^2},\quad \eta=-x-\sqrt{x^2+y^2}.$$

这两个值 ξ,η 可以看做是 xy 平面上的曲线坐标, 而共焦点抛物线就是它的坐标曲线. 这些都已在图 3.9 中表示出来, 只要我们想像着把符号 (x,y) 与 (ξ,η) 互相调换一下就成了.

在应用抛物线坐标 (ξ,η) 时, 我们必须记住一对值 (ξ,η) 对应着**两个**点 (x,y) 与 $(x,-y)$, 即相应的抛物线上的两个交点. 因此, 为要得到数对 (x,y) 到数对 (ξ,η) 间的一一对应, 我们就必须限制 (x,y) 于半个平面, 譬如 $y\geq 0$, 这样, 就在这个半平面上每个区域 R 都与 $\xi\eta$ 平面上的一个区域 B 一一对应, 而这个区域 B 上的每

一个点的直角坐标 (ξ,η) 就与区域 R 内的相应点的抛物线坐标完全相同了.

练 习 3.3 b

1. 证明, 对于 $x \neq 1, 0 < y < \pi/2$, 方程组

$$\xi = (\sin y)/(x-1), \quad \eta = x\tan y$$

定义了一个曲线坐标系.

2. 求圆周 $x^2+y^2=1$ 在曲线坐标系

$$\xi = x^3+1, \quad \eta = xy$$

下的方程.

3. 在 (x,y) 平面上的哪些点不能用

$$\xi = xy, \quad \eta = x^2+y^2$$

作为曲线坐标?

c. 推广到多于两个变量的情形

对于三个或更多自变量, 情况完全同上面类似. 例如一组定义在 (x,y,z) 空间的一个区域 R 内的三个连续可微的函数

$$\xi = \phi(x,y,z), \quad \eta = \psi(x,y,z), \quad \zeta = \chi(x,y,z)$$

可以看作区域 R 到 (ξ,η,ζ) 空间的一个区域 B 的一个映射. 如果这个由 R 到 B 的映射是 1 对 1 的, 因而对于 B 的每一个像点 (ξ,η,ζ) 而言 R 内的相应点 (即原点或逆像) 的坐标 (x,y,z) 都可以由函数组

$$x = g(\xi,\eta,\zeta), \quad y = h(\xi,\eta,\zeta), \quad z = l(\xi,\eta,\zeta)$$

唯一地计算出来，那么 (ξ,η,ζ) 就可以看做区域 R 内的点 P 的一种广义坐标. 曲面 $\xi=$ 常数，$\eta=$ 常数，$\zeta=$ 常数，或者改用另一种写法，

$$\phi(x,y,z)=\text{常数},\quad \psi(x,y,z)=\text{常数},\quad \chi(x,y,z)=\text{常数},$$

于是形成由覆盖区域 R 的三族曲面组成的曲面组，可以称之为**曲线坐标面**.

完全同两个变量的情形一样，我们可以把三维空间中的**一一变换**理解为连续分布在空间的某整个区域上的介质的变形.

一个非常重要的坐标系是**球坐标**，有时也称为**空间的极坐标**. 这种坐标系用这样三个数来描绘空间中的一点 P 的位置：(1) 它与原点的距离

$$r=\sqrt{x^2+y^2+z^2};$$

(2) 地理经度 ϕ，即 xz 平面与由 P 点和 z 轴所确定的平面之间的夹角；(3) 极倾角或叫余纬 θ，即向径 OP 与正 z 轴之间的夹角. 从图 3.10 可以看到，三个球坐标 r,ϕ,θ 与直角坐标是由这样的变换方程联系着的：

$$\begin{aligned}
x&=r\cos\phi\sin\theta,\\
y&=r\sin\phi\sin\theta,\\
z&=r\cos\theta;
\end{aligned}$$

由此可以推得其逆变换关系式为

$$\begin{aligned}
&r=\sqrt{x^2+y^2+z^2},\\
&\phi=\arccos\frac{x}{\sqrt{x^2+y^2}}=\arcsin\frac{y}{\sqrt{x^2+y^2}},\\
&\theta=\arccos\theta\frac{z}{\sqrt{x^2+y^2+z^2}}=\arcsin\frac{\sqrt{x^2+y^2}}{\sqrt{x^2+y^2+z^2}}.
\end{aligned}$$

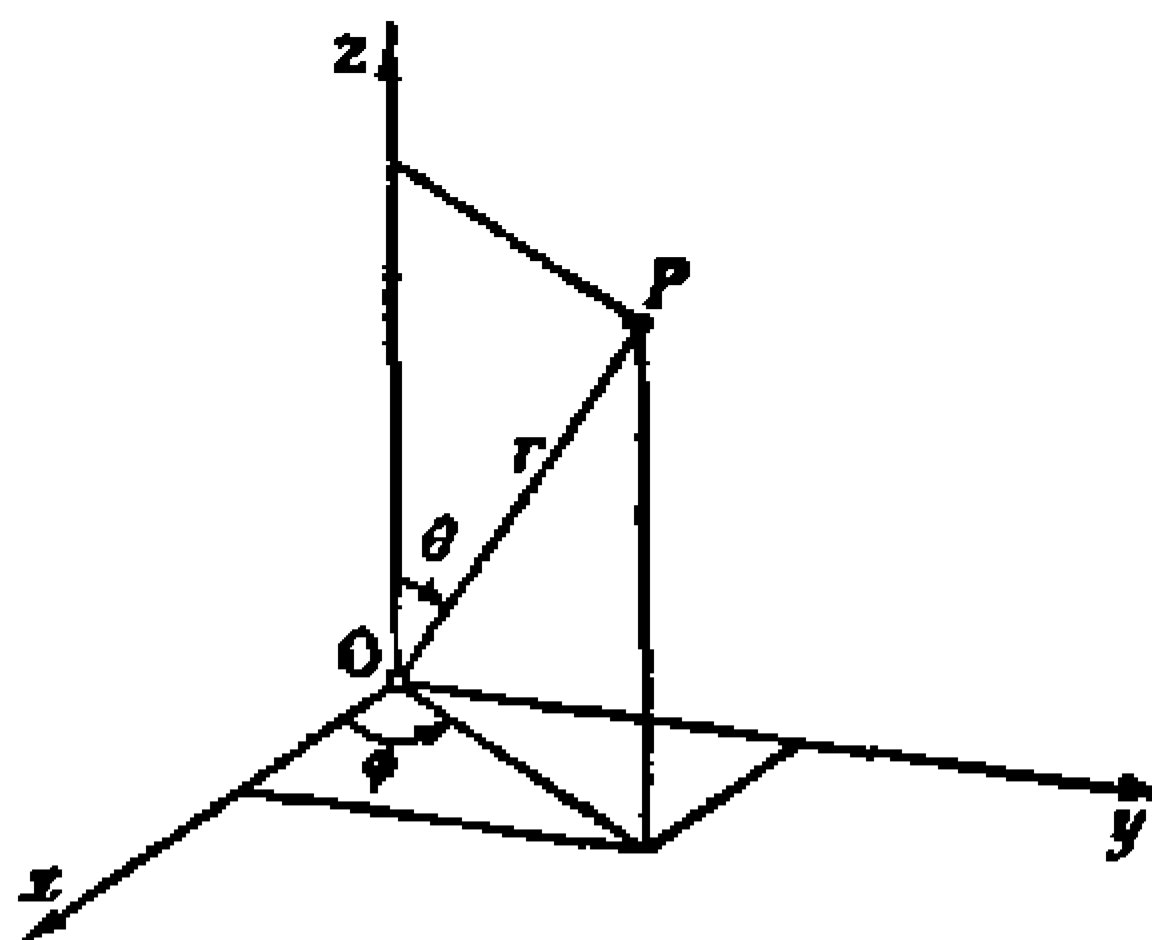


图 3.10　球坐标

在平面的极坐标中原点是 $1-1$ 变换失效的唯一的例外点，因为唯有原点的幅角是不定的．同样地在空间的球坐标下，整个 z 轴上的点由于经度不定而都是例外；至于原点本身，则极倾角 θ 也是不定的．

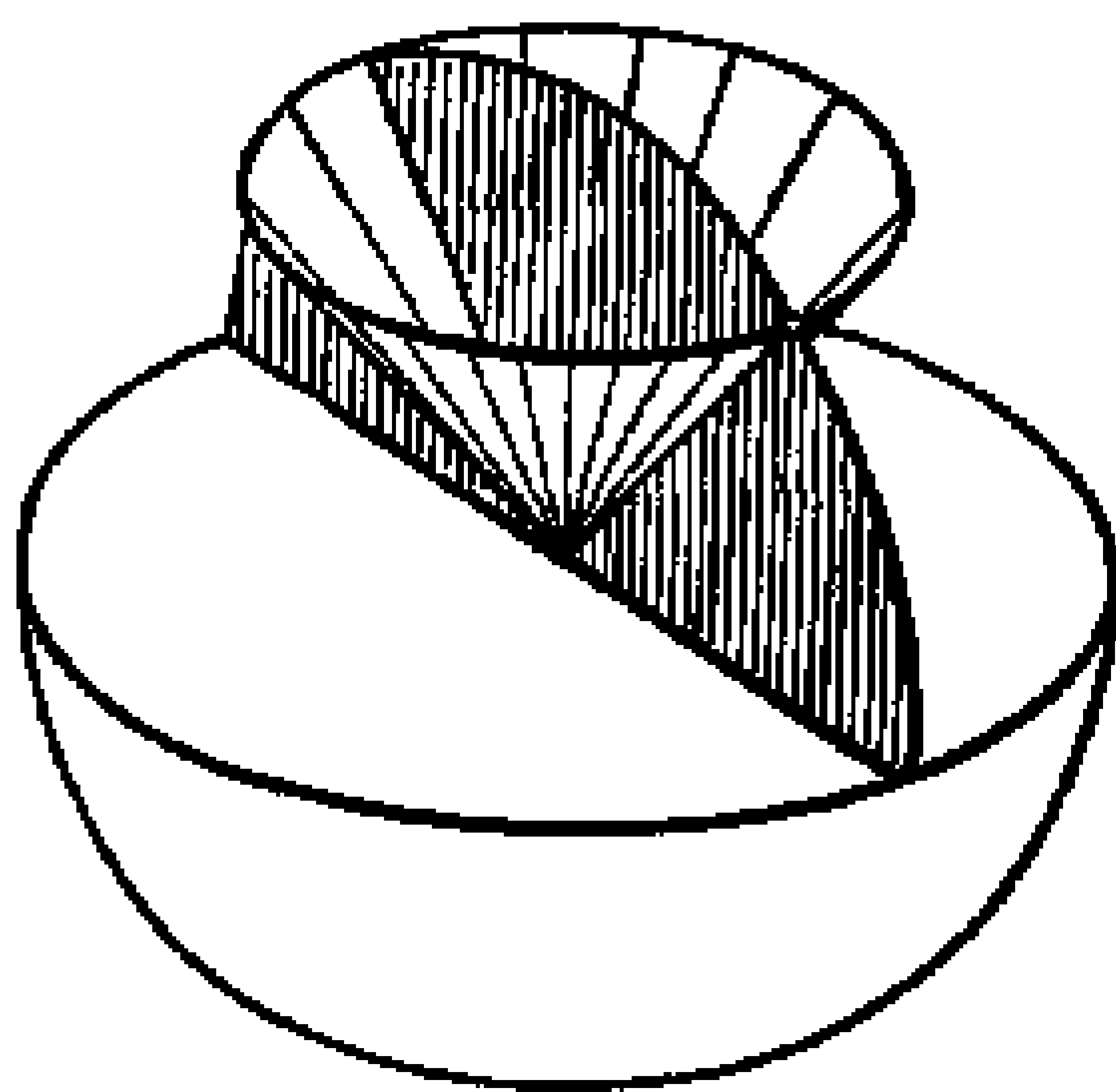


图 3.11　球坐标的坐标面

三维极坐标的坐标面如下：(1) 对 r 的固定值，是关于原点的同心球；(2) 对 ϕ 的固定值，是过 z 轴的半平面族；(3) 对 θ 的固定值，是以 z 轴为轴而以原点为顶点的一族圆锥面 (图 3.11)．

另一种常用的坐标系是 **柱坐标**．这是由直角坐标系在 xy 平面上改用极坐标系而保留 z 作为第三个坐标得到的．从而由直角坐标

到柱坐标的变换公式是

$$x=\rho\cos\phi,$$

$$y=\rho\sin\phi,$$

$$z=z;$$

而其逆变换公式则是

$$\rho=\sqrt{x^2+y^2},$$

$$\phi=\arccos\frac{x}{\sqrt{x^2+y^2}}=\arcsin\frac{y}{\sqrt{x^2+y^2}},$$

$$z=z.$$

坐标面 $\rho=$ 常数是一族正圆柱，它们与 xy 平面相交于以原点为心的同心圆，而坐标面 $\phi=$ 常数则是过 z 轴的半平面族；而坐标面 $z=$ 常数则是与 xy 平面平行的平面族.

练　习　3.3 c

1. 求曲线坐标变换

$$\xi=\frac{x}{x^2+y^2+z^2},\quad \eta=\frac{y}{x^2+y^2+z^2},\quad \zeta=\frac{z}{x^2+y^2+z^2}$$

的逆变换.

2. 求坐标变换

$$w=r\cos\phi,$$

$$x=r\sin\phi\cos\psi,$$

$$y=r\sin\phi\sin\psi\cos\theta,$$

$$z=r\sin\phi\sin\psi\sin\theta$$

的逆变换．并问：$r=$ 常数，$\phi=$ 常数，$\psi=$ 常数，$\theta=$ 常数，各代表什么样的集合？

d. 反函数的微商公式

许多很重要的情形中，像以上的例子一样，可以把函数组解成显式，从而可知反函数是否连续与是否有连续微商．但是，如果事先我们知道反函数的存在性与可微性，我们就可以用下述方式来计算反函数的微商，而无需把方程实际按显式解出：我们把反函数 $x=g(\xi,\eta)$ $y=h(\xi,\eta)$ 代入方程 $\xi=\phi(x,y)$, $\eta=\psi(x,y)$, 在右边就得到 ξ,η 的复合函数 $\phi(g(\xi,\eta),\ h(\xi,\eta))$ 与 $\psi(g(\xi,\eta),h(\xi,\eta))$, 而它们一定分别等于 ξ,η. 我们现在可以对下列每一个方程关于 ξ,η 分别求微商

$$\xi=\phi(g(\xi,\eta),\ h(\xi,\eta))\quad \eta=\psi(g(\xi,\eta),\ h(\xi,\eta)),\tag{24a}$$

即把 ξ,η 看成自变量[1]，并用链式法则对复合函数求微商．这样就得到方程组

$$1=\phi_x g_\xi+\phi_y h_\xi,\quad 0=\phi_x g_\eta+\phi_y h_\eta,$$

$$0=\psi_x g_\xi+\psi_y h_\xi,\quad 1=\psi_x g_\eta+\psi_y h_\eta.\tag{24b}$$

解这些方程便得到反函数 $x=g(\xi,\eta)$, $y=h(\xi,\eta)$ 关于 ξ,η 的偏微商，它们表示成为原来函数 $\phi(x,y)$, $\psi(x,y)$ 关于 x,y 的偏微商形式便是

$$g_\xi=\frac{\psi_y}{D},\quad g_\eta=\frac{-\phi_y}{D},\quad h_\xi=\frac{-\psi_x}{D},\quad h_\eta=\frac{\phi_x}{D},\tag{24c}$$

或写成

$$x_\xi=\frac{\eta_y}{D},\quad x_\eta=\frac{-\xi_y}{D},\quad y_\xi=\frac{-\eta_x}{D},\quad y_\eta=\frac{\xi_x}{D}.\tag{24d}$$

---

1) 这些方程对所考虑的 ξ,η 的一切值都成立，即它们是*恒等地*成立的，不同于只对这些变量中的*某些值*才成立的方程组．这种恒等方程或*恒等式*，当对出现在方程中的任何变量进行微商时，由定义可知仍然给出恒等式．

这里 D 是下式的缩写

$$D = \xi_x\eta_y - \xi_y\eta_x = \begin{vmatrix} \dfrac{\partial\xi}{\partial x} & \dfrac{\partial\xi}{\partial y} \\ \dfrac{\partial\eta}{\partial x} & \dfrac{\partial\eta}{\partial y} \end{vmatrix}. \tag{24e}$$

表达式 D(我们假定它在所考虑的点不等于 0) 叫做函数 $\xi = \phi(x,y)$, $\eta = \psi(x,y)$, 关于变量 x,y 的 **雅可比**(Jacobi)**行列式** 或 **函数行列式**. 它在我们考虑变换时总是扮演着很重要的角色, 这一点在后面的讨论中将变得越来越明显.

以上 (以及偶尔还在别处) 我们使用了较简短的符号 $\xi(x,y)$ 代替那较为细致的符号 $\xi = \phi(x,y)$, 后者区分了量 ξ 及其函数表达式 $\phi(x,y)$. 以后, 如无混淆的情况, 我们将常用类似的简化记号.

对于把平面的极坐标表示成直角坐标的表示式,

$$\xi = r = \sqrt{x^2+y^2} \quad 与 \quad \eta = \theta = \operatorname{arctg}\frac{y}{x}, \tag{25}$$

偏微商为

$$r_x = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{r}, \quad r_y = \frac{y}{\sqrt{x^2+y^2}} = \frac{y}{r},$$

$$\theta_x = \frac{-y}{x^2+y^2} = -\frac{y}{r^2}, \quad \theta_y = \frac{x}{x^2+y^2} = \frac{x}{r^2}.$$

因此, 雅可比行列式取值为

$$D = \frac{x}{r}\cdot\frac{x}{r^2} - \frac{y}{r}\left(-\frac{y}{r^2}\right) = \frac{1}{r},$$

而反函数的偏微商 (把直角坐标表示成极坐标) 则根据 (24d) 应是

$$x_r = \frac{x}{r}, \quad x_\theta = -y, \quad y_r = \frac{y}{r}, \quad y_\theta = x.$$

这当然可以更容易地从反函数公式 $x = r\cos\theta$, $y = r\sin\theta$ 直接求微商得出来.

由于雅可比行列式经常用到，所以常给以一特殊符号 [1]

$$D=\frac{d(\xi,\eta)}{d(x,y)}.$$

这种缩写是否合适，我们马上就会明白. 由反函数的微商公式 (24d) 我们可以得到函数 $x=x(\xi,\eta)$ 与 $y=y(\xi,\eta)$ 关于 ξ,η 的雅可比行列式，它由下列式子给出

$$\frac{d(x,y)}{d(\xi,\eta)}=x_\xi y_\eta-x_\eta y_\xi=\frac{\xi_x\eta_y-\xi_y\eta_x}{D^2}$$

$$=\frac{1}{D}=\left(\frac{d(\xi,\eta)}{d(x,y)}\right)^{-1}. \tag{26}$$

这就是说，**反函数组的雅可比行列式是原函数组的雅可比行列式的倒数** [2].

我们也可以把反函数组的二阶微商用原函数组的一阶、二阶微商表示出来. 我们只需对 ξ 与 η 用链式法则微分线性方程 (24b). (当然，我们要假定所给函数具有二阶连续微商.) 于是，我们得到线性方程，从而很容易计算出所需要的微商.

例如，要计算二阶微商

$$\frac{\partial^2 x}{\partial\xi^2}=g_{\xi\xi}\text{与}\ \frac{\partial^2 y}{\partial\xi^2}=h_{\xi\xi},$$

我们再一次用链式法则对 ξ,η 微分这两个方程

$$1=\xi_x x_\xi+\xi_y y_\xi,$$

$$0=\eta_x x_\xi+\eta_y y_\xi,$$

---

1) 常常把雅可比行列式用偏微商符号写出

$$D=\frac{\partial(\xi,\eta)}{\partial(x,y)}.$$

2) 这与一元函数的反函数微商法则很相似 (见第一卷第 233 页).

便得到

$$0 = \xi_{xx}x_\xi^2 + 2\xi_{xy}x_\xi y_\xi + \xi_{yy}y_\xi^2 + \xi_x x_{\xi\xi} + \xi_y y_{\xi\xi}, \qquad (27a)$$

$$0 = \eta_{xx}x_\xi^2 + 2\eta_{xy}x_\xi y_\xi + \eta_{yy}y_\xi^2 + \eta_x x_{\xi\xi} + \xi_y y_{\xi\xi}. \qquad (27b)$$

如果我们把 $x_{\xi\xi}, y_{\xi\xi}$ 看成未知数来解这个线性方程组 (这方程组行列式还是 D, 因此按假定也不是 0) 并用已知值 x_ξ, y_ξ 来代替, 经过简短计算得到

$$x_{\xi\xi} = -\frac{1}{D^3}\begin{vmatrix} \xi_{xx}\eta_y^2 - 2\xi_{xy}\eta_x\eta_y + \xi_{yy}\eta_x^2 & \xi_y \\ \eta_{xx}\eta_y^2 - 2\xi_{xy}\eta_x\eta_y + \eta_{yy}\eta_x^2 & \eta_y \end{vmatrix}, \qquad (27c)$$

$$y_{\xi\xi} = -\frac{1}{D^3}\begin{vmatrix} \xi_{xx}\eta_y^2 - 2\xi_{xy}\eta_x\eta_y + \xi_{yy}\eta_x^2 & \xi_x \\ \eta_{xx}\eta_y^2 - 2\xi_{xy}\eta_x\eta_y + \eta_{yy}\eta_x^2 & \eta_x \end{vmatrix}. \qquad (27d)$$

三阶以及更高阶都可由同样方法即反复微分这线性方程组得到, 而每一步我们都得到线性方程组并有非 0 的行列式 D.

练　习　3.3 d

1. 求下列变换的雅可比行列式:

(a) $\xi = ax + by, \eta = cx + dy$;

(b) $r = \sqrt{x^2 + y^2}, \theta = \operatorname{arctg} y/x$;

(c) $\xi = x^2, \eta = y^2$;

(d) $\xi = \frac{1}{2}\log(x^2 + y^2), \eta = \operatorname{arctg} y/x$;

(e) $\xi = xy^2, \eta = x^2y$;

(f) $\xi = x^3 - y, \eta = y^3 + x$.

2. 对习题 1 中所给的每一个变换, 找出这样的点 (x, y), 它们没有任何邻域在其中存在逆变换.

3. 对下列的每一情况, 求出变换 $\xi = f(x, y), \eta = g(x, y)$ 的雅可比行列式, 以及 x, y 关于 ξ, η 的直到二阶的全部偏导数:

(a) $\xi = e^x \cos y$, $\eta = e^x \sin y$;

(b) $\xi = x^2 - y^2$, $\eta = 2xy$;

(c) $\xi = \operatorname{tg}(x+y), \eta = \cos(x-y), -\dfrac{\pi}{2} < x + y < \dfrac{\pi}{2}$;

(d) $\xi = \sinh x + \cosh y$, $\eta = -\cosh x + \sinh y$;

(e) $\xi = x^3 + y^3$, $\eta = xy^2$.

4. 称一个变换为 "保角的"(见第 316 页), 如果在该变换下任何两条相应曲线的交角保持不变.

(a) 证明反演变换

$$\xi = \frac{x}{x^2+y^2}, \qquad \eta = \frac{y}{x^2+y^2}$$

是保角的;

(b) 证明任一圆的反演像是另一圆或直线;

(c) 求出反演变换的雅可比行列式.

5. 设 K_1, K_2, K_3 是经过原点的三个圆, 并有相异的交点 P_1, P_2, P_3. 证明由圆弧所组成的曲线三角形 $P_1P_2P_3$ 的三角之和等于 π.

6. 平面的一个变换

$$u = \varphi(x,y), \quad v = \psi(x,y)$$

是保角的, 如果函数 φ 与 ψ 满足恒等式

$$\varphi_x = \psi_y; \qquad \varphi_y = -\psi_x.$$

7. 证明, 若一曲面 $z = u(x,y)$ 的所有法线都与 z 轴相交, 则这曲面是一个旋转面.

8. 方程

$$\frac{x^2}{a-t} + \frac{y^2}{b-t} = 1 \qquad (a > b)$$

决定 t 的依赖于 x, y 的两个值:

$$t_1 = \lambda(x,y), \quad t_2 = \mu(x,y).$$

(a) 证明曲线 $t_1=$ 常数与 $t_2=$ 常数是具有相同焦点的椭圆与双曲线 (共焦点的圆锥曲线).

(b) 证明曲线 $t_1=$ 常数, $t_2=$ 常数是互相正交的.

(c) t_1 与 t_2 可用作曲线坐标 (称为焦点坐标). 试用这些坐标来表示 x 和 y.

(d) 试用 x 与 y 来表示雅可比行列式 $\partial(t_1,t_2)/\partial(x,y)$.

(e) 试求出在焦点坐标系下由参量方程

$$t_1 = f_1(\lambda),\quad t_2 = f_2(\lambda) \quad 与 \quad t_1 = g_1(\mu),\quad t_2 = g_2(\mu)$$

表示的两族曲线相互正交的条件.

9. (a) 证明 t 的方程

$$\frac{x^2}{a-t}+\frac{y^2}{b-t}+\frac{z^2}{c-t}=1 \quad (a>b>c)$$

有三个不同的实根 t_1,t_2,t_3, 分别属于区间

$$-\infty<t<c,\quad c<t<b,\quad b<t<a,$$

这里假定了点 (x,y,z) 不在坐标面上.

(b) 证明通过一任意点的三个曲面 $t_1=$ 常数, $t_2=$ 常数, $t_3=$ 常数是相互正交的.

(c) 把 x,y,z 用焦点坐标 t_1,t_2,t_3 表示出来.

10. 证明, 由方程

$$\xi=\frac{1}{2}\left(x+\frac{x}{x^2+y^2}\right),\quad \eta=\frac{1}{2}\left(y-\frac{y}{x^2+y^2}\right)$$

给出的 xy 平面上的变换

(a) 是保角的;

(b) 把 (x,y) 平面上过原点的直线与以原点为圆心的圆变换成由方程

$$\frac{\xi^2}{t+\frac{1}{2}}+\frac{\eta^2}{t-\frac{1}{2}}=1$$

给出的共焦点的圆锥曲线 $t=$ 常数.

11. 对于 $\xi=f(x,y)$, $\eta=g(x,y)$ 与

$$D=\partial(\xi,\eta)/\partial(x,y)\neq 0,$$

证明下列恒等式:

(a) $\dfrac{\partial D}{\partial y}=\dfrac{\partial(\xi_y,\eta)}{\partial(x,y)}+\dfrac{\partial(\xi,\eta_y)}{\partial(x,y)}$,

(b) $D^{-3}[\xi_x(\eta_{yy}D-\eta_yD_y)-\xi_y(\eta_{xy}D-\eta_yD_x)]$
$\quad =D^{-3}[\eta_x(\xi_{yy}D-\xi_yD_y)-\eta_y(\xi_{xy}D-\xi_yD_x)]$.

e. 映射的符号乘积

我们从关于变换的复合作一些注记开始. 设变换

$$\xi=\phi(x,y),\quad \eta=\psi(x,y) \tag{28a}$$

给出由区域 R 的点 (x,y) 到 $\xi\eta$ 平面上的区域 B 的点 (ξ,η) 的一个一一映射; 又设方程

$$u=\Phi(\xi,\eta),\quad v=\Psi(\xi,\eta) \tag{28b}$$

给出由区域 B 到 uv 平面上的区域 R' 的一个一一映射. 于是产生一个由 R 到 R' 的一一映射. 这个映射很自然地称为**结果映射**或**变换**, 并且说是由该两个给定的映射复合得到的, 成为它们的**符号乘积**. 这个结果变换按定义是由方程组

$$u=\Phi(\phi(x,y),\psi(x,y)),\quad v=\Psi(\phi(x,y),\psi(x,y))$$

给出的, 从而可以立即推知这个映射也是一一的.

由复合函数的微商法则, 我们得到

$$\frac{\partial u}{\partial x}=\Phi_\xi\phi_x+\Phi_\eta\psi_x,\quad \frac{\partial u}{\partial y}=\Phi_\xi\phi_y+\Phi_\eta\psi_y, \tag{29a}$$

$$\frac{\partial v}{\partial x} = \overline{\Psi}_\xi \phi_x + \overline{\Psi} \psi_{\eta x}, \quad \frac{\partial v}{\partial y} = \overline{\Psi}_\xi \phi_y + \overline{\Psi}_\eta \psi_y. \tag{29b}$$

用矩阵的符号 (第 164 页) 表示就是

$$\begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} \Phi_\xi & \Phi_\eta \\ \overline{\Psi}_\xi & \overline{\Psi}_\eta \end{pmatrix} \begin{pmatrix} \phi_x & \phi_y \\ \psi_x & \psi_y \end{pmatrix}. \tag{30}$$

把这个同行列式乘积法则 (见第 186 页) 相比较, 我们就发现 [1] u, v 关于 x, y 的雅可比行列式是

$$\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x} = (\Phi_\xi \overline{\Psi}_\eta - \Phi_\eta \overline{\Psi}_\xi)(\phi_x \psi_y - \phi_y \psi_x). \tag{31a}$$

用话来说, **两个变换的符号乘积的雅可比行列式正好等于各个变换的雅可比行列式的乘积**

$$\frac{d(u,v)}{d(x,y)} = \frac{d(u,v)}{d(\xi,\eta)} \cdot \frac{d(\xi,\eta)}{d(x,y)}. \tag{31b}$$

这个方程正好显示出我们用以表示雅可比行列式的符号是恰当的. **当变换被复合时, 其雅可比行列式的关系就同一元函数被复合时微商之间的关系一样**. 当分别变换的雅可比行列式都不为 0 时, 结果变换的雅可比是不会为 0 的.

特别地, 如果第二个变换

$$u = \phi(\xi, \eta), \quad v = \psi(\xi, \eta)$$

是第一个的逆

$$\xi = \phi(x, y), \quad \eta = \psi(x, y),$$

而两个变换又都是可微的, 则结果变换将只不过是恒等变换, 即 $u = x$, $v = y$; 这最后一个雅可比行列式显然是 1; 从而我们又得到关系 (26).

---

1) 当然, 由直接乘积也可得到这结果.

由此推知两个雅可比行列式中的任何一个都不等于 0, 并且

$$\frac{d(\xi,\eta)}{d(x,y)}\cdot\frac{d(x,y)}{d(\xi,\eta)}=1.$$

对于一对连续可微的函数组 $\phi(x,y)$ 与 $\psi(x,y)$, 若有非 0 的雅可比行列式, 我们就能够找到相应的映射在一点 $P_0=(x_0,y_0)$ 处的方向的公式. 一条经过点 P_0 的曲线可以由参数方程 $x=f(t)$, $y=g(t)$ 来描绘, 其中 $f(t_0)=x_0$, $g(t_0)=y_0$. 曲线在 P_0 的斜率是

$$m=\frac{g'(t_0)}{f'(t_0)};$$

而像曲线

$$\xi=\varphi(f(t),g(t)),\quad \eta=\psi(f(t),g(t))$$

在相应于 P_0 的那一点的斜率是

$$\mu=\frac{d\eta/dt}{d\xi/dt}=\frac{\psi_x f'+\psi_y g'}{\phi_x f'+\phi_y g'}=\frac{c+dm}{a+bm},\tag{32}$$

其中 a,b,c,d 是常数:

$$a=\phi_x(x_0,y_0),\quad b=\phi_y(x_0,y_0),$$

$$c=\psi_x(x_0,y_0),\quad d=\psi_y(x_0,y_0).$$

原来曲线在 P_0 的斜率 m 与像曲线的斜率 μ 之间的关系 (32) 就与在 P_0 的邻域内近似于映射的仿射映射

$$\xi=\phi(x_0y_0)+a(x-x_0)+b(y-y_0),$$

$$\eta=\psi(x_0y_0)+c(x-x_0)+d(y-y_0).$$

的关系相同. 因为

$$\frac{d\mu}{dm}=\frac{ad-bc}{(a+bm)^2},$$

因而我们看到，当 $ad-bc>0$ 时 μ 是 m 的一个递增函数，当

$$ad-bc<0$$

时 μ 是 m 的一个递减函数[1]).

斜率的增加相应于倾斜角的增加，也就是相应方向的逆时针转动. 因此，$d\mu/dm>0$ 时隐含着逆时针的旋转方向保持不变；而 $d\mu/dm<0$ 时，则反过来. 现在，$ad-bc$ 却正好就是雅可比行列式

$$\frac{d(\xi,\eta)}{d(x,y)}=\begin{vmatrix}\phi_x & \phi_y\\ \psi_x & \psi_y\end{vmatrix}$$

在点 P_0 处计算出的值，于是推知，**映射** $\xi=\phi(x,y),\eta=\psi(x,y)$ **在** (x_0y_0) **附近是保持还是改变旋转定向，要看雅可比行列式在该点的值是正的还是负的.**

练　习　3.3 e

1. 对于下列每一对变换用两种方法求 $\partial(u,v)/\partial(x,y)$，先消去 ξ,η 来计算，然后用公式 (31b) 来计算：

(a) $\begin{cases} u=\dfrac{1}{2}\log(\xi^2+\eta^2),\\ v=\operatorname{arctg}\eta/\xi;\end{cases}$ $\begin{cases}\xi=e^x\cos y,\\ \eta=e^x\sin y.\end{cases}$

(b) $\begin{cases} u=\xi^2-\eta^2,\\ v=2\xi\eta;\end{cases}$ $\begin{cases}\xi=x\cos y,\\ \eta=x\sin y.\end{cases}$

(c) $\begin{cases} u=e^\xi\cos\eta,\\ v=e^\xi\sin\eta;\end{cases}$ $\begin{cases}\xi=x/(x^2+y^2),\\ \eta=-y/(x^2+y^2).\end{cases}$

2. 在下列的复合变换中哪些能够把 x,y 定义成 u,v 在所指定

1) 更确切地说，这只是在局部上成立，要除去 m 或 μ 变成无穷这样的方面.

点 (u_0, v_0) 的邻域内的连续可微函数?

(a) $\xi = e^x \cos y$, $\eta = e^x \sin y$; $u = \xi^2 - \eta^2$, $v = 2\xi\eta$, $u_0 = 1$, $v_0 = 0$;

(b) $\xi = \cosh x + \sinh y$, $\eta = \sinh x + \cosh y$; $u = e^{\xi+\eta}$, $v = e^{\xi-\eta}$, $u_0 = v_0 = 1$;

(c) $\xi = x^3 - y^3$, $\eta = x^2 + 2xy^2$; $u = \xi^5 + \eta$, $v = \eta^5 - \xi$; $u_0 = 1$, $v_0 = 0$.

3. 考虑变换 $\begin{cases} u = \phi(\xi, \eta), \\ v = \psi(\xi, \eta), \end{cases}$ $\begin{cases} \xi = f(x), \\ \eta = g(y). \end{cases}$

证明

$$\frac{\partial(u, v)}{\partial(x, y)} = f'(x) \cdot g'(y) \frac{\partial(u, v)}{\partial(\xi, \eta)}.$$

4. 设 $z = f(x, y)$ 与 $\xi = \phi(x, y)$, $\eta = \psi(x, y)$, 并假定 $\dfrac{\partial(\xi, \eta)}{\partial(x, y)} \neq 0$, 试验证

$$\frac{\partial z}{\partial \xi} = \frac{\partial(z, \eta)}{\partial(x, y)} \Big/ \frac{\partial(\xi, \eta)}{\partial(x, y)},$$

$$\frac{\partial z}{\partial \eta} = \frac{\partial(\xi, z)}{\partial(x, y)} \Big/ \frac{\partial(\xi, \eta)}{\partial(x, y)}.$$

f. 关于变换及隐函数组的逆的一般定理. 分解成素映射

变换求逆的可能性依赖于下述的一般定理:

设 $\phi(x, y)$ 和 $\psi(x, y)$ 在点 (x_0, y_0) 的邻域内是连续可微的函数, 其雅可比行列式 $D = \phi_x\psi_y - \phi_y\psi_x$ 在点 (x_0, y_0) 处不为零. 命 $u_0 = \phi(x_0, y_0)$, $v_0 = \psi(x_0, y_0)$. 则存在点 (x_0, y_0) 的一个邻域 N 和点 (u_0, v_0) 的一个邻域 N', 使得映射

$$u = \phi(x, y), \quad v = \psi(x, y) \tag{33a}$$

有一个唯一的逆映射

$$x = g(u, v), \quad y = h(u, v) \tag{33b}$$

将 N' **映入** N. **对** N' **内的** (u,v), **函数** g **和** h **满足恒等式**

$$u=\phi(g(u,v),\ h(u,v)),\ v=\psi(g(u,v),h(u,v)), \tag{33c}$$

并满足方程

$$x_0=g(u_0,v_0),\quad y_0=h(u_0,v_0). \tag{33d}$$

反函数 g,h **在** (u_0,v_0) **附近对** (u,v) **有连续的微商，它们由下式给出**

$$\frac{\partial x}{\partial u}=\frac{1}{D}\frac{\partial v}{\partial y},\quad \frac{\partial x}{\partial v}=-\frac{1}{D}\frac{\partial u}{\partial y}, \tag{33e}$$

$$\frac{\partial y}{\partial u}=-\frac{1}{D}\frac{\partial v}{\partial x},\quad \frac{\partial y}{\partial v}=\frac{1}{D}\frac{\partial u}{\partial x}. \tag{33f}$$

证明可从第 248 页上的隐函数定理推出，该定理允许人们从一个方程中解出一个变量来. 本质上，我们求方程组 (33a) 的逆的办法是，从第一个方程中解出变量 x,y 中的一个，并将所得的表达式代入第二个方程，得出一个只含第二个变量的方程.

因为依假设，雅可比行列式 D 在点 (x_0,y_0) 不为零，所以在该点至少有 $\phi(x,y)$ 的一个一阶导数异于零，譬如说是 $\phi_x(x_0,y_0)\neq 0$. 这时我们就能关于 x 解方程

$$u=\phi(x,y). \tag{34a}$$

说得更确切些，我们能找到正的常数 h_1,h_2,h_3，使得对

$$|u-u_0|<h_1,\quad |y-y_0|<h_2 \tag{34b}$$

方程 (34a) 有唯一解 $x=X(u,y)$ 满足 $|x-x_0|<h_3$. 函数 $X(u,y)$ 有定义区域 (34b) 并满足方程

$$\phi(X(u,y),y)=u,\quad X(u_0,y_0)=x_0, \tag{34c}$$

和不等式

$$|X(u,y)-x_0|<h_3. \tag{34d}$$

此外， $X(u,y)$ 有连续的微商，由于 (34c) 而满足

$$\phi_x(X(u,y),y)\,X_u(u,y)=1, \tag{34e}$$

$$\phi_x(X(u,y),y)X_y(u,y)+\phi_y(X(u,y),y)=0. \tag{34f}$$

这里我们假定 h_2, h_3 很小，使矩形

$$|x-x_0|<h_3,\quad |y-y_0|<h_2 \tag{34g}$$

落在 $\phi(x,y)$; $\psi(x,y)$ 的定义区域中. 以表达式 $X(u,y)$ 代函数 $\psi(x,y)$ 中的 x, 我们得到复合函数

$$\psi(X(u,y),y)=\chi(u,y), \tag{34h}$$

它有定义域 (34b). 这里，由 (34c, f),

$$\chi(u_0,y_0)=\psi(x_0,y_0)=v_0, \tag{34i}$$

$$\begin{aligned}\chi_y(u_0,y_0)&=\psi_x x_y+\psi_y=-\psi_x\frac{\phi_y}{\phi_x}+\psi_y\\&=\frac{D}{\phi_x}\neq 0;\end{aligned} \tag{34j}$$

由 (34e) 知， $\phi_x\neq 0$. 于是我们能找到正的常数 h_4,h_5,h_6, 使得对于

$$|u-u_0|<h_4,\quad |v-v_0|<h_5, \tag{34k}$$

方程

$$\chi(u,y)=v \tag{34l}$$

有唯一解 $y=h(u,v)$, 它满足 $|y-y_0|<h_6$. 这里我们可以假定 $h_4\leq h_1$, $h_6\leq h_2$(见第 248 页脚注).

最后，我们令

$$X(u,h(u,v))=g(u,v). \tag{34m}$$

两个函数 $g(u,v)$, $h(u,v)$ 都有定义域 (34k). 由 (34c, h) 它们满足方程

$$\phi(g(u,v),h(u,v)) = \phi(X(u,h(u,v)),h(u,v)) = u,$$

$$\psi(g(u,v),h(u,v)) = \psi(X(u,h(u,v)),h(u,v))$$

$$= \chi(u,h(u,v)) = v,$$

和不等式

$$|g(u,v) - x_0| < h_3, |h(u,v) - y_0| < h_6.$$

g 和 h 的微商公式曾更早地在第 247 页上导出过.

为了证明反函数的唯一性, 假定 x, y, u, v 是满足方程 (33a) 和不等式

$$|x - x_0| < h_3, |y - y_0| < h_6, \ |u - u_0| < h_4, \ |v - v_0| < h_5$$

的任何一组值. 因 (34a, b) 成立, 我们断定

$$x = X(u,y). \tag{34n}$$

从 (34h) 我们得到方程

$$v = \psi(x,y) = \psi(X(u,y),y) = \chi(u,y),$$

这个方程有唯一解 $y = h(u,v)$. 于是由 (34n, o) 推出关系式 $x = g(u,v)$. 关于 g 和 h 的关系式 (33d), 从解的唯一性和假定 $u_0 = \phi(x_0,y_0)$, $v_0 = \psi(x_0,y_0)$ 便可推出.

到此为止, 我们一直假定 $\phi_x(x_0,y_0) \neq 0$. 如果 $\phi_x(x_0,y_0) = 0$, 而 $\phi_y(x_0,y_0) \neq 0$, 求映射 (33a) 的逆映射的步骤是类似的. 在这种情况下, 我们从 (33a) 的第一个方程解出 y, 把所得函数 $y = Y(u,x)$ 代入第二个方程, 便得到一个只含 x 的方程.

平面映射 (33a) 的求逆问题曾经 (如上) 被归结成了暂时只有一个变量作变换的那种映射的求逆问题. 一般地说, 如果变换 (33a)

容许一个坐标不变化，即如果或者函数 $\phi(x,y)$ 恒等于 x，或者函数 $\psi(x,y)$ 恒等于 y，则我们称它为一个**素变换**．形如 $u=\phi(x,y)$，$v=y$ 的素变换的效果是沿 x 轴的方向移动每一个点，而保持其纵坐标不变．在变换之后这个点具有新的横坐标，它既依赖于 x，也依赖于 y．如果这素变换的雅可比式 ϕ 是正的，则对固定的 y，u 随 x 单调递增地变化．

我们将证明，**在一个点的邻域内，任何一个具有非零的雅可比式的变换**(33a)**都能分解成素变换之积**．这个事实容易从我们对逆映射的构造中推出．如果 $\phi_x(x_0,y_0)\neq 0$，我们就把映射 (33a) 表为素映射

$$\xi=\phi(x,y),\quad \eta=y \tag{34o}$$

和

$$u=\xi,\quad v=x(\xi,\eta) \tag{34p}$$

的符号乘积．这里，第一映射在 xy 平面上的区域 R 是一个矩形，它很小，以至

$$|x-x_0|<h_3,\ |y-y_0|<h_2,\ |\phi(x,y)-u_0|<h_1;$$

而第二映射的区域是

$$|\xi-u_0|<h_1,\quad |\eta-y_0|<h_2.$$

于是，在映射 (34p) 之下，R 的每个点 (x,y) 的像 (ξ,η) 总是落在映射 (34q) 的区域中，并且

$$x=X(\xi,y).$$

因此，又有

$$x=X(\phi(x,y),y). \tag{34q}$$

这时，根据 (34h,r)，对于由 (34p,q) 复合而成的映射，我们有

$$u=\phi(x,y),$$

$$v=\chi(\phi(x,y),y)=\psi(X(\phi(x,y),y),y)=\psi(x,y).$$

当 $\phi_x(x_0,y_0)=0$ 而 $\phi_y(x_0,y_0)\neq 0$ 时，我们只要改换 x 和 y 在变换中的地位，就得到映射 (33a) 的一个类似的分解.

我们不能期望在整个开区域 R 内都是以同一种方式将一个变换分解为素变换. 但是，由于在 R 的每一个点的邻近可以用某种方式把变换进行分解，所以 R 的每个有界闭子集可被分为有限个闭子集 [1]，使变换在每个闭子集上总有某一种方式的分解是可能的.

逆变换定理仅是一个更一般的定理的一种特殊情况，该一般定理可视为隐函数定理对于函数组的一种推广. 隐函数定理 (第 239 页) 适合于对一个方程解出其中一个变量. 一般定理如下：

如果 $\phi(x,y,u,v,\cdots,w)$ **和** $\psi(x,y,u,v,\cdots,w)$ **是** $x,y,u,v,\cdots,w$ **的连续可微函数，并且存在某一点** $(x_0,y_0,u_0,v_0,\cdots,w_0)$ **满足方程组**

$$\phi(x,y,u,v,\cdots,w)=0 \text{ 和 } \psi(x,y,u,v,\cdots,w)=0,$$

又如果在该点 ϕ **和** ψ **关于** x **和** y **的雅可比式不为零**(即，$D=\phi_x\psi_y-\phi_y\psi_x\neq 0$)**则在该点的邻域内，方程** $\phi=0$ **和** $\psi=0$ **能且仅能以一种方式解出** x **和** y，**并且** x **和** y **是** $u,v,\cdots,w$ **的连续可微函数.**

这个定理的证明类似于上面的逆变换定理的证明. 从假设 $D\neq 0$，我们能够推断，在所考虑的那个点有某个偏微商不为零，譬如说 $\phi_x\neq 0$. 根据第 248 页上的主要定理，如果我们把 $x,y,u,v,\cdots,w$ 分别限制在 $x_0,y_0,u_0,v_0,\cdots,w_0$ 附近的足够小的区间内，则方程 $\phi(x,y,u,v,\cdots,w)=0$ 恰能以一种方式将 x 解为其他变量的函数，而且这个解 $x=X(y,u,v,\cdots,w)$ 是自变量的连续可微函数，并具有偏微商 $X_y=-\phi_y/\phi_x$. 如果我们把函数 $x=X(y,u,v,\cdots,w)$ 代入 $\psi(x,y,u,v,\cdots,w)$，我们就得到一个函数 $\psi(x,y,u,v,\cdots,w)=$

1) 这从第 118 页上的覆盖定理推出.

$\chi(y,u,v,\cdots,w)$, 并且

$$\chi_y = -\psi_x\frac{\phi_y}{\phi_x} + \psi_y = \frac{D}{\phi_x}.$$

依假设 $D\neq 0$, 因此, 我们知道偏微商 χ_y 是不为零的. 于是, 如果我们把 $y,u,v,\cdots,w$ 限制在 $y_0,u_0,v_0,\cdots,w_0$ 附近的区间中, 并使这些区间落在上面预先限定的区间中, 则我们仅能以一种方式将方程 $\chi=0$ 解为 $u,v,\cdots,w$ 的函数 y, 并且这个解是连续可微的. 把 y 的这个表达式代入方程 $\chi=X(y,u,v,\cdots,w)$ 中, 我们就找到了函数 x, 它作为 $u,v,\cdots,w$ 的函数. 这个解是唯一的并且是连续可微的, 其中 $x,y,u,v,\cdots,w$ 被分别限制在 $x_0,y_0,u_0,v_0,\cdots,w_0$ 附近的足够小的区间内.

练 习 3.3 f

1. 在指定点附近, 下列方程组的哪一个能将 x,y 解为其他变量的连续可微函数?

(a) $\begin{cases} e^x\sin u - e^y\cos v + w = 0, \\ x\cosh w - u\sinh y - v^2 = \cosh 1, \end{cases}$
$x=1,\ y=0,\ u=0,\ v=0,\ w=1;$

(b) $\begin{cases} u\cos x - v\sin y + w^2 = 1, \\ \cos(x+y) + v = 1, \end{cases}$
$x=0, y=\dfrac{\pi}{2}, u=1, v=1, w=1;$

(c) $\begin{cases} x^2+y^2+u^2-v=0, \\ x^2-y^2+2u-1=0, \end{cases}$
$x=y=u=v=1;$

(d) $\begin{cases} \cos x + t\sin y = 0, \\ \sin x - \cos ty = 0, \end{cases}$
$x=\pi,\ y=\dfrac{\pi}{2},\ t=1.$

g. 用逐次逼近法迭代构造逆映射

在前面的证明中, 求逆映射的问题被归结为一维情况, 而最后又归结为一个基本事实: 由一个变量的连续单调函数所提供的映射

是可逆的. 这种讨论有两点缺陷: 我们不得不区分导致结果十分不同的不同情况 (譬如说, 对 $\phi_x \neq 0$ 和 $\phi_x = 0$), 而其不同结果并不相应于原来的变换在特性上的任何根本变化; 而且, 证明是存在性的而 非构造性的, 对求逆映射它并没有提供一个实际的数值计算方法. 这两个缺陷在迭代法或逐次逼近法中都不出现. 逐次逼近法出自第一卷, 它是为了求解含一个未知量的方程而给出的数值方法的一种范式. 其基本思想是对近似解作逐次修正, 在这里, 修正解是由在一点邻域内对函数关系作最佳逼近的 线性方程 所确定的.

我们仍考虑方程

$$u = \varphi(x, y), \quad v = \psi(x, y), \tag{35a}$$

其中 φ 和 ψ 是 xy 平面上的开集 R 内的连续可微函数. 设 (x_0, y_0) 是 R 内的一点, 在这一点雅可比式

$$\begin{vmatrix} \varphi_x & \varphi_y \\ \psi_x & \psi_y \end{vmatrix} \tag{35b}$$

的值不为零. 设在映射 (35a) 之下, (u_0, v_0) 是 (x_0, y_0) 的像点. 我们要证明的是对足够接近 (u_0, v_0) 的 (u, v), 存在着接近于 (x_0, y_0) 的唯一确定的值 (x, y), 满足 $u = \varphi(x, y)$ 和 $v = \psi(x, y)$.

为了求得解, 我们将运用和第一卷中对一个变量函数的讨论一样的迭代法, 而使用一种适合于二维情况的记号. 引进向量 $\mathbf{U} = (u, v)$, $\mathbf{X} = (x, y)$, 我们能将映射 (35a) 简单地写成形式

$$\mathbf{U} = F(\mathbf{X}), \tag{35c}$$

这里 F 是一个非线性变换, 它把以 x, y 为分量的向量映射为以 $\phi(x, y)$, $\psi(x, y)$ 为分量的向量. 微分 dx, dy 和 du, dv 满足线性关系 (见第 61 页)

$$du = d\varphi = \varphi_x dx + \varphi_y dy, \tag{35d}$$

$$dv = d\psi = \psi_x dx + \psi_y dy. \tag{35e}$$

如果把微分法合成向量写成 $d\mathbf{X} = (dx, dy)$, $d\mathbf{U} = (du, dv)$, 我们就能将关系式 (35d,e) 写成 [1]

$$d\mathbf{U} = F' d\mathbf{X}, \tag{35f}$$

其中 $\mathbf{F}'$ 是由映射函数的一阶微商形成的二阶矩阵

$$\mathbf{F}' = \begin{pmatrix} \varphi_x & \varphi_y \\ \psi_x & \psi_y \end{pmatrix}. \tag{35g}$$

显然, 矩阵 $\mathbf{F}'$ 起着向量映射函数 $\mathbf{F}$ 的微商作用. $\mathbf{F}'$ 的行列式恰是映射的雅可比式 (35b)[2]. 一般地, 我们为了强调矩阵 $\mathbf{F}'$ 对向量 $\mathbf{X} = (x, y)$ 的依赖关系而写作 $\mathbf{F}' = \mathbf{F}'(\mathbf{X})$. 对线性变换而言, 矩阵 $\mathbf{F}'$ 是常数.

矩阵 $\mathbf{F}'$ 的元素的 "大小" 限制了映射 $\mathbf{F}$ 能把距离放大多少. 取两点 (x, y) 和 $(x + h, y + k)$, 使连接它们的直线段整个位于映射的区域之内. 根据多元函数的中值定理 (第 73 页),

$$\begin{aligned} \varphi(x + h, y + k) - \varphi(x, y) &= \varphi'_x h + \varphi_y k, \\ \psi(x + h, y + k) - \psi(x, y) &= \psi_x h + \psi_y k, \end{aligned} \tag{36}$$

其中一阶微商在连接 (x, y) 和 $(x + h, y + k)$ 的线段上的适当点处取值 [3]. 设 M 是诸量

$$|\varphi_x|, |\varphi_y|, |\psi_x|, |\psi_y|$$

---

1) 最好将 (35f) 解释为三个矩阵 $d\mathbf{U}, \mathbf{F}', d\mathbf{X}$ 之间的一种关系, 其中 $d\mathbf{X}, d\mathbf{U}$ 等同于二行一列的矩阵:

$$d\mathbf{X} = \begin{pmatrix} dx \\ dy \end{pmatrix} \quad d\mathbf{U} = \begin{pmatrix} du \\ dv \end{pmatrix};$$

见第 159 页.

2) 矩阵 $\mathbf{F}'$ 常称做雅可比矩阵或映射的弗雷西微商.

3) 一般说来, 第一个方程的中间点和第二个方程的中间点是不同的.

在连接 (x,y) 和 $(x+h,y+k)$ 的线段上的全部点所取值的一个上界. 那么, 显然, 像点间的距离可估计如下:

$$\begin{aligned}&\sqrt{(\varphi(x+h,y+k)-\varphi(x,y))^2+(\psi(x+h,y+k)-\psi(x,y))^2}\\ \leq&\sqrt{(M|h|+M|k|)^2+(M|h|+M|k|)^2}\\ =&\sqrt{2}M(|h|+|k|)\leq 2M\sqrt{h^2+k^2}.\end{aligned}\tag{36a}$$

因此, 像点间的距离至多是原来两点间距离的 $2M$ 倍. 引进向量 $\mathbf{Y}=(x+h,y+k)$, 我们就能将 (36a) 写成关于映射 $\mathbf{F}$ 的利普希茨条件的形式

$$|\mathbf{F}(\mathbf{Y})-\mathbf{F}(\mathbf{X})|\leq 2M|\mathbf{Y}-\mathbf{X}|,\tag{36b}$$

其中 M 是矩阵 $\mathbf{F}'$ 的诸元素的绝对值的一个上界 [1]. 用矩阵记号, (36) 变为

$$\mathbf{F}(\mathbf{Y})-\mathbf{F}(\mathbf{X})=\mathbf{H}(\mathbf{X},\mathbf{Y})(\mathbf{Y}-\mathbf{X}),\tag{36c}$$

这里矩阵 $\mathbf{H}$ 满足

$$\lim_{Y\to X}\mathbf{H}(\mathbf{X},\mathbf{Y})=\mathbf{F}'(\mathbf{X}).\tag{36d}$$

今在 $\mathbf{F}$ 的区域 R 内的点 $X_0=(x_0,y_0)$ 的邻域

$$|\mathbf{X}-\mathbf{X}_0|<\delta\tag{37a}$$

内来考虑映射 $\mathbf{U}=\mathbf{F}(\mathbf{X})$. 设 $\mathbf{U}_0=\mathbf{F}(\mathbf{X}_0)=(u_0,v_0)$. 对固定的 $\mathbf{U}$, 我们把要解 $\mathbf{X}$ 的方程 $\mathbf{U}=\mathbf{F}(\mathbf{X})$ 写成如下形式

$$\mathbf{X}=\mathbf{G}(\mathbf{X}),\tag{37b}$$

其中

$$\mathbf{G}(\mathbf{X})=\mathbf{X}+\mathbf{a}(\mathbf{U}-\mathbf{F}(\mathbf{X}));\tag{37c}$$

---

1) 在映射 $\mathbf{F}$ 是 n 维的情况, (36b) 中的因子 2 要用 n 来代替.

这里 $\mathbf{a}$ 是一个适当选取的非奇异的常数矩阵，它有逆 $\mathbf{a}^{-1}$. 这时方程 (37b) 等价于 $\mathbf{a}(\mathbf{U}-\mathbf{F}(\mathbf{X}))=0$, 用 $\mathbf{a}^{-1}$ 乘它得到

$$\mathbf{a}^{-1}\mathbf{a}(\mathbf{U}-\mathbf{F}(\mathbf{X}))=\mathbf{e}(\mathbf{U}-\mathbf{F}(\mathbf{X}))=\mathbf{U}-\mathbf{F}(\mathbf{X})=\mathbf{0},$$

其中 $\mathbf{e}$ 是单位矩阵. 于是, (37b) 的任何一个解 $\mathbf{X}$—— 也就是映射 $\mathbf{G}$ 的任何一个不动点—— 提供 $\mathbf{U}=\mathbf{F}(\mathbf{X})$ 的一个解.

我们将证明, (37b) 的一个解 $\mathbf{X}$ 可由 $\mathbf{X}_n$ 的极限给出, $\mathbf{X}_n$ 由递推公式

$$\mathbf{X}_{n+1}=\mathbf{G}(\mathbf{X}_n) \quad (n=0,1,2,\cdots) \tag{37d}$$

确定; 条件是要表示向量映射 $\mathbf{G}$ 的微商的矩阵 $\mathbf{G}'(\mathbf{X})$ 足够小. 更精确些说, 对 $\mathbf{X}_0$ 的邻域 (37a) 内的所有 $\mathbf{X}$, 我们要求, 矩阵 $\mathbf{G}'$ 的元素的最大绝对值比 1/4 小, 并且

$$|G(X_0)-X_0|<\frac{1}{2}\delta.$$

首先, 我们用归纳法证明, 在所述的假定之下, 递推公式 (37d) 只引出满足 (37a) 的向量. 这就肯定了 $\mathbf{X}_n$ 位于 G 的区域之内, 从而序列能无限地持续下去. 当 $M=1/4$ 时, 由 (36b) 我们发现,

$$|\mathbf{G}(\mathbf{Y})-\mathbf{G}(\mathbf{X})|\le\frac{1}{2}|\mathbf{Y}-\mathbf{X}|$$

对

$$|\mathbf{X}-\mathbf{X}_0|<\delta, \quad |\mathbf{Y}-\mathbf{X}_0|<\delta. \tag{37e}$$

现在, 不等式 (37a) 对 $\mathbf{X}=\mathbf{X}_0$ 是显然满足的. 如果它对 $\mathbf{X}=\mathbf{X}_n$ 成立, 则对由 (37d) 所确定的向量 $\mathbf{X}_{n+1}$, 我们得到,

$$\begin{aligned}|\mathbf{X}_{n+1}-\mathbf{X}_0|&\le|\mathbf{X}_{n+1}-\mathbf{X}_1|+|\mathbf{X}_1-\mathbf{X}_0|\\&=|\mathbf{G}(\mathbf{X}_n)-\mathbf{G}(\mathbf{X}_0)|+|\mathbf{G}(\mathbf{X}_0)-\mathbf{X}_0|\\&\le\frac{1}{2}|\mathbf{X}_n-\mathbf{X}_0|+\frac{1}{2}\delta<\delta.\end{aligned}$$

这就证明了，对于所有的 n 都有 $|\mathbf{X}_n-\mathbf{X}_0|<\delta$.

为了看出 $\mathbf{X}_n$ 收敛，根据 (37e) 我们注意到

$$
\begin{aligned}
|\mathbf{X}_{n+1}-\mathbf{X}_n| &= |\mathbf{G}(\mathbf{X}_n)-\mathbf{G}(\mathbf{X}_{n-1})| \\
&\le \frac{1}{2}|\mathbf{X}_n-\mathbf{X}_{n-1}|.
\end{aligned}
$$

同理

$$
|\mathbf{X}_n-\mathbf{X}_{n-1}| \le \frac{1}{2}|\mathbf{X}_{n-1}-\mathbf{X}_{n-2}|,
$$

$$
|\mathbf{X}_{n-1}-\mathbf{X}_{n-2}| \le \frac{1}{2}|\mathbf{X}_{n-2}-\mathbf{X}_{n-3}|,
$$

如此下去，这些不等式一起引出估计

$$
|\mathbf{X}_{n+1}-\mathbf{X}_n| \le \frac{1}{2^n}|\mathbf{X}_1-\mathbf{X}_0| \le \frac{\delta}{2^{n+1}}. \tag{37f}
$$

把 $\mathbf{X}$ 写作无穷级数之和

$$
\begin{aligned}
\mathbf{X} = \mathbf{X}_0 + (\mathbf{X}_1-\mathbf{X}_0) + (\mathbf{X}_2-\mathbf{X}_1) + \cdots \\
+ (\mathbf{X}_{n+1}-\mathbf{X}_n) + \cdots,
\end{aligned}
$$

其收敛性是根据 (37f) 把它与收敛的几何级数 相比较 (第一卷第 584 页) 而建立的；由此也就推出了 $\mathbf{X}=\lim \mathbf{X}_n$ 的存在性. 当 $n\to\infty$ 时，用 $\mathbf{G}(\mathbf{X})$ 的连续性，从 (37d) 立刻就可推出 $\mathbf{X}$ 是 (37b) 的一个解.

根据定义 (37c)，函数 G 不仅连续地依赖于 $\mathbf{X}$，而且也连续地依赖于向量 $\mathbf{U}$. 因此，由递推公式 (37d) 逐次得出的 $\mathbf{X}_n$ 也连续地依赖于 $\mathbf{U}$[1]. 由于用作比较的几何级数在建立

$$
\mathbf{X} = \lim_{n\to\infty} \mathbf{X}_n
$$

1) 这里我们用了连续函数的连续函数仍是连续函数这一事实.

时不依赖于 $\mathbf{U}$, 这就推出 $\mathbf{X}$ 是 $\mathbf{U}$ 的连续函数的一致极限, 因此 $\mathbf{X}$ 本身就是 $\mathbf{U}$ 的一个连续函数. 并且显然, 由于对所有的 n,

$$|\mathbf{X}_n - \mathbf{X}| < \delta,$$

所以 $|\mathbf{X} - \mathbf{X}_0| \leq \delta$. 要是存在满足条件 $\mathbf{Y} = \mathbf{G}(\mathbf{Y})$ 和 $|\mathbf{Y} - \mathbf{X}_0| \leq \delta$ 的第二个解 $\mathbf{Y}$, 由 (37e) 我们就会发现

$$|\mathbf{Y} - \mathbf{X}| = |\mathbf{G}(\mathbf{Y}) - \mathbf{G}(\mathbf{X})| \leq \frac{1}{2}|\mathbf{Y} - \mathbf{X}|,$$

因此 $|\mathbf{Y} - \mathbf{X}| = 0$, 从而 $\mathbf{Y} = \mathbf{X}$.

这样, 对 $|\mathbf{X} - \mathbf{X}_0| \leq \delta$, 我们确立了方程 $\mathbf{U} = \mathbf{F}(\mathbf{X})$ 的解 $\mathbf{X}$ 的存在性, 唯一性和连续性. 条件是, 由 (37c) 所确定的向量 $\mathbf{G}$ 有微商 $\mathbf{G}'$, 而对 $|\mathbf{X} - \mathbf{X}_0| \leq \delta$, $\mathbf{G}'$ 的元素的绝对值小于 1/4, 而且还有条件

$$|\mathbf{G}(\mathbf{X}_0) - \mathbf{X}_0| < \frac{1}{2}\delta.$$

易见, 适当选取矩阵 $\mathbf{a}$, 对所有充分接近 $\mathbf{U}_0$ 的 $\mathbf{U}$ 这些要求是能够满足的. 根据 (37c),

$$\mathbf{G}'(\mathbf{X}) = \mathbf{e} - \mathbf{a}\mathbf{F}'(\mathbf{X}),$$

这里 $\mathbf{e}$ 是单位矩阵. 如果我们将 $\mathbf{a}$ 选为矩阵 $\mathbf{F}'(\mathbf{X}_0)$ 的逆矩阵,

$$\mathbf{a} = (\mathbf{F}'(\mathbf{X}_0))^{-1},$$

则对 $\mathbf{X} = \mathbf{X}_0$ 有

$$\mathbf{G}'(\mathbf{X}_0) = \mathbf{e} - \mathbf{a}\mathbf{F}'(\mathbf{X}_0) = \mathbf{0}.$$

(从矩阵 $\mathbf{F}'(\mathbf{X}_0)$ 有非零行列式, 也就是映射 $\mathbf{F}$ 的雅可比式在点 $\mathbf{X}_0$ 不为零这一基本假定便能推导出这个逆矩阵的存在性.) 从假定映射 $\mathbf{F}$ 的一阶微商的连续性导出 $\mathbf{G}'(\mathbf{X})$ 连续地依赖于 $\mathbf{X}$; 因此, 对足够小的 $|\mathbf{X} - \mathbf{X}_0|$, 譬如说

$$|\mathbf{X} - \mathbf{X}_0| \leq \delta,$$

便能保证 $G'(X)$ 的元素是任意小，例如比 1/4 小；此外，根据 (37c)，

$$|\mathbf{G}(\mathbf{X}_0)-\mathbf{X}_0|=|\mathbf{a}(\mathbf{U}-\mathbf{F}(\mathbf{X}_0))|=|\mathbf{a}(\mathbf{U}-\mathbf{U}_0)|<\frac{1}{2}\delta,$$

只要 $\mathbf{U}$ 位于 $\mathbf{U}_0$ 的足够小的邻域内即可.

这就完成了具有非零雅可比式的连续可微映射的连续逆映射的局部存在性的证明. 从 (36c, d) 容易推出逆映射的一阶微商的存在性和连续性. 设 $\mathbf{U}=\mathbf{F}(\mathbf{X})$，这里我们假定雅可比矩阵 $\mathbf{F}'(\mathbf{X})$ 是非奇异的. 则每一个足够接近 $\mathbf{U}$ 的 $\mathbf{V}$ 具有形式 $\mathbf{V}=\mathbf{F}(\mathbf{Y})$，其中 $\mathbf{Y}$ 当 $\mathbf{V}$ 趋近于 $\mathbf{U}$ 时趋近于 $\mathbf{X}$. 因此，对足够接近于 $\mathbf{U}$ 的 $\mathbf{V}$ 矩阵 $\mathbf{H}(\mathbf{X},\mathbf{Y})$ 也是非奇异的. 这时我们发现

$$\begin{aligned}\mathbf{Y}-\mathbf{X}&=(\mathbf{H}(\mathbf{X},\mathbf{Y}))^{-1}(\mathbf{V}-\mathbf{U})\\&=(\mathbf{F}'(\mathbf{X}))^{-1}(\mathbf{V}-\mathbf{U})+\mathbf{E}(\mathbf{X},\mathbf{Y})(\mathbf{V}-\mathbf{U}),\end{aligned}$$

其中

$$\lim_{V\to U}\mathbf{E}(\mathbf{X},\mathbf{Y})=\lim_{Y\to X}\mathbf{E}(\mathbf{X},\mathbf{Y})=0.$$

但是，这个关系式恰好表示了满足 $\mathbf{U}=\mathbf{F}(\mathbf{X})$ 的向量 $\mathbf{X}$ 是向量 $\mathbf{U}$ 的可微函数，并且，$\mathbf{X}$ 关于 $\mathbf{U}$ 的雅可比矩阵是矩阵 $\mathbf{F}'(\mathbf{X})$ 的逆. 显然，用**迭代法或逐次逼近法**构造逆映射的方法同样能运用到任意维数的映射上去.

练　习　3.3 g

1. 对于

$$u=\frac{1}{2}(x^2-y^2),\quad v=xy$$

的逆映射在 $\mathbf{X}=(1,1)$ 或 $\mathbf{U}=(0,1)$ 的邻域内运用 (37d) 求出迭代近似解 (x_2,y_2).

2. 将前一练习的结果与 $u=1, v=1$ 的邻域内 x 和 y 的泰勒展式的二阶项相比较.

b. 函数的相依性

如果在点 (x_0, y_0) 雅可比式 D 为零，则在该点的邻域内，关于方程 (33a) 的可解性问题不再有一个一般的论断了．即使逆函数组确实存在，它们也不可微了，因为这时乘积

$$\frac{d(u,v)}{d(x,y)} \cdot \frac{d(x,y)}{d(u,v)}$$

将等于零，而根据第 263 页，它必须等于 1．例如，方程组

$$u = x^3, \quad v = y$$

虽然雅可比式在原点为零，仍唯一可解，其解为

$$x = \sqrt[3]{u}, \quad y = v;$$

但函数 $\sqrt[3]{u}$ 在原点不可微．

另一方面，方程组

$$u = x^2 - y^2, \quad v = 2xy$$

在原点的邻域内不是唯一可解，因为 xy 平面上的两点 (x, y) 和 $(-x, -y)$ 都对应于 uv 平面上的同一个点．

如果雅可比式不仅在单独一个点 (x, y) 为零，而且在点 (x, y) 的某一整个邻域内恒为零，则变换称为**退化的**．在这种情况下，函数

$$u = \varphi(x, y) \quad \text{和} \quad v = \psi(x, y)$$

呈现一种相依性，即它们中有一个是另外一个的函数[1]．我们首先考虑方程 $\varphi_x = 0$ 和 $\varphi_y = 0$ 处处成立的平凡情况，这时函数 $\varphi(x, y)$ 是一个常数．由此我们就看到，当点 (x, y) 在整个区域上变化时，

---

1) 雅可比式为零也等价于由映射函数的一阶微商所构成的**向量** (φ_x, φ_y) 和 (ψ_x, ψ_y) **相依**．

它的像点 (u,v) 总保持在直线 $u=$ 常数上. 这就是说, 区域只是被映射到一条直线上而不是一个区域上. 所以这时不可能在两个二维区域之间存在一个到另一个上的一一映射.

微商 φ_x 和 φ_y 至少有一个不为零而雅可比式 D 仍为零的一般情况也类似. 假定在所考虑的区域的点 (x_0,y_0) 处我们有 $\varphi_x \neq 0$. 于是从第一个方程中可解出 x, 形如 $x = X(u,y)$, 并像第 267 页那样写成 $v = \psi(X(u,y),y) = \chi(u,y)$, 因为在那里我们只用了 $\varphi_x \neq 0$ 的假定. 但是根据 (34j) 和方程 $D = 0$, 在区域中 $\varphi_x \neq 0$ 的地方必有 χ_y 等于零; 这就是说, 量 $\chi = v$ 根本不依赖于 y, 而 v 仅是 u 的函数. 于是我们得到结论: 如果变换的雅可比式恒为零, 则在此变换下, xy 平面的一个区域不再映射为 uv 平面上的一个区域, 而是映射为 uv 平面上的一条曲线, 因为在值 u 的某个区间内每个 u 值仅有一个值 v 与之对应. 因此, 如果雅可比式恒为零, 则两函数不独立; 这就是说存在一个关系式

$$F(\varphi,\psi) = \psi - \chi(\varphi) = 0,$$

使得在区域中的一切值 (x,y) 都满足它. 反之, 如果在 uv 平面上存在一条曲线, xy 平面上的区域映射于其上, 则对于这个区域上的一切点雅可比式 $D = \varphi_x\psi_y - \varphi_y\psi_x$ 必恒为零, 因为, 显然这映射在一点的一整个邻域内是不可逆的.

在开始各别讨论的特殊情况显然包含在一般的论断之中. 那时问题中的曲线恰是直线 $u =$ 常数, 它平行于 v 轴.

例如

$$\xi = x + y, \quad \eta = (x+y)^2$$

是一个退化的变换. 这个变换把 xy 平面上的一切点都映射为 $\xi\eta$ 平面上的抛物线 $\eta = \xi^2$ 的点. 变换求逆的问题是谈不到的, 因为直线 $x+y=$ 常数上的所有点都被映射成了一个点 (ξ,η). 容易验证, 雅可比式的值是 0. 与一般定理一致, 函数 ξ 和 η 之间的关系由方程

$$F(\xi,\eta) = \xi^2 - \eta = 0$$

给出.

练 习 3.3 h

1. 给出一对连续可微函数 $\xi = f(x,y)$, $\eta = g(x,y)$ 的例子，使它们在一个区域内是独立的，而在另一个区域内是不独立的.

2. 证明，如果 $\xi = ax + by + c$ 和 $\eta = \alpha x + \beta y + \gamma$ 是相依的，则直线 $\xi = 0$ 和 $\eta = 0$ 是平行的.

i. 结束语

将上述理论推广到三个或三个以上自变量的情况不再出现特殊的困难. 主要的差别在于，我们用三阶或多阶行列式代替二阶行列式 D. 当变换含有三个自变量时，

$$\xi = \varphi(x,y,z), \quad \eta = \psi(x,y,z), \quad \zeta = \chi(x,y,z),$$

$$x = g(\xi,\eta,\zeta), \quad y = h(\xi,\eta,\zeta), \quad z = l(\xi,\eta,\zeta),$$

雅可比式由下式给出:

$$D = \frac{d(\xi,\eta,\zeta)}{d(x,y,z)} = \begin{vmatrix} \varphi_x & \psi_x & \chi_x \\ \varphi_y & \psi_y & \chi_y \\ \varphi_z & \psi_z & \chi_z \end{vmatrix}. \tag{38}$$

同样，对于 n 个自变量的变换

$$\xi_i = \varphi_i(x_1, x_2, \cdots, x_n)$$

$$x_i = g_i(\xi_1, \xi_2, \cdots, \xi_n) \quad (i = 1, 2, \cdots, n)$$

而言，雅可比式为

$$\frac{d(\xi_1,\xi_2,\cdots,\xi_n)}{d(x_1,x_2,\cdots,x_n)}=\begin{vmatrix}\frac{\partial\varphi_1}{\partial x_1} & \frac{\partial\varphi_2}{\partial x_1} & \cdots & \frac{\partial\varphi_n}{\partial x_1}\\ \frac{\partial\varphi_1}{\partial x_2} & \frac{\partial\varphi_2}{\partial x_2} & \cdots & \frac{\partial\varphi_n}{\partial x_2}\\ \vdots & \vdots & & \vdots\\ \frac{\partial\varphi_1}{\partial x_n} & \frac{\partial\varphi_2}{\partial x_n} & \cdots & \frac{\partial\varphi_n}{\partial x_n}\end{vmatrix}.$$

当自变量多于两个时，复合变换的雅可比式等于各个雅可比式的乘积，这一命题仍然对. 用符号写就是

$$\frac{d(\xi_1,\xi_2,\cdots,\xi_n)}{d(\eta_1,\eta_2,\cdots,\eta_n)}\cdot\frac{d(\eta_1,\eta_2,\cdots,\eta_n)}{d(x_1,x_2,\cdots,x_n)}$$
$$=\frac{d(\xi_1,\xi_2,\cdots,\xi_n)}{d(x_1,x_2,\cdots,x_n)}.$$

特别地，逆变换的雅可比式是原变换雅可比式的倒数.

关于变换的分解与合成的定理，关于逆变换的定理，关于变换相依性的定理，对于自变量是三个或更多的情况仍然成立. 其证明类似于 $n=2$ 的情况，为了避免不必要的重复我们就略去了. 用迭代法构造逆映射也同样成立.

在前一节我们看到了，一个一般变换在许多方面类似于仿射变换，而雅可比式起着在仿射变换中行列式所起的作用. 下面的说明使这点更加清楚. 由于函数 $\xi=\varphi(x,y)$ 和 $\eta=\psi(x,y)$ 在 (x_0,y_0) 的邻域内是可微的，我们可将它们表示为如下形式：

$$\begin{aligned}\xi-\xi_0=&(x-x_0)\varphi_x(x_0,y_0)+(y-y_0)\varphi_y(x_0,y_0)\\&+\varepsilon\sqrt{(x-x_0)^2+(y-y_0)^2},\\ \eta-\eta_0=&(x-x_0)\psi_x(x_0,y_0)+(y-y_0)\psi_y(x_0,y_0)\\&+\delta\sqrt{(x-x_0)^2+(y-y_0)^2},\end{aligned}$$

其中 ε 和 δ 随着

$$\sqrt{(x-x_0)^2+(y-y_0)^2}$$

趋向于零而趋向于零. 这表明, 当 $|x-x_0|$ 和 $|y-y_0|$ 的值足够小时, 变换能被近似地表成仿射变换:

$$\xi=\xi_0+(x-x_0)\varphi_x(x_0,y_0)+(y-y_0)\varphi_y(x_0,y_0),$$

$$\eta=\eta_0+(x-x_0)\psi_x(x_0,y_0)+(y-y_0)\psi_y(x_0,y_0),$$

它的行列式就是原变换的雅可比式.

练 习 3.3 i

1. 对下列函数组计算 $\partial(\xi,\eta,\rho)/\partial(x,y,z)$:

(a) $\xi=e^x\cos y\cos z,$

$\eta=e^x\cos y\sin z,$

$\rho=e^x\sin y;$

(b) $\xi=\cos(x+y)+\cos(y+z),$

$\eta=\cos(x+y)+\sin(y+z),$

$\rho=\sin(x+y)+\cos(y+z);$

(c) $\xi=\cosh x+\log y,$

$\eta=\tanh y-\sinh z,$

$\rho=x-y^z;$

(d) $\xi=x\cos y\sin z,$

$\eta=x\sin y\sin z,$

$\rho=x\cos z;$

(e) $\xi=x\cos y,$

$\eta=x\sin y,$

$\rho=z.$

2. 定义函数组 $\xi = f(x,y,z)$, $\eta = g(x,y,z)$, $\rho = h(x,y,z)$ 在一区域内的相依性. 将 h 段的结果推广到这种情况.

3. 在习题 1 中所给的三个函数所成的组中, 那一组是相依的? 给出三个函数间的联系方程.

4. 证明下列三个函数是相依的, 并找出它们之间的相依关系:

$$\xi = x + y + z,$$

$$\eta = x^2 + y^2 + z^2,$$

$$\zeta = xy + yz + zx.$$

5. 三维反演由公式

$$\xi = \frac{x}{x^2+y^2+z^2}, \quad \eta = \frac{y}{x^2+y^2+z^2},$$

$$\zeta = \frac{z}{x^2+y^2+z^2}$$

定义.

(a) 证明任何两曲面间夹角是不变的;

(b) 证明球面被变换为球面或平面;

(c) 求变换的雅可比式.

3.4 应　　用

a. 曲面理论的要素

曲面像曲线一样, 参变量表示法常常是比其他表示法更合用. 对曲面我们需要两个参变量, 而不是一个. 我们用 u 和 v 来表示它们. 曲面的参变量表示可表成如下形式:

$$x = \varphi(u,v), \quad y = \psi(u,v), \quad z = \chi(u,v), \tag{39a}$$

其中 φ, ψ 和 χ 都是参量 u,v 的已知函数, 而点 (u,v) 在 uv 平面上的已给区域 R 上变化. 从而以 x,y,z 为三维直角坐标的相应点

就在 xyz 空间的一个集合上变化. 在典型情况下, 这集合是一张曲面, 能表为显式 $z=f(x,y)$. 因为对上列三个方程, 可将其中两个根据相应的直角坐标解出 u,v, 然后把所得到的 u 和 v 的表达式代入第三个方程, 我们就得到了曲面的非对称表示 $z=f(x,y)$[1]. 因此, 为了保证方程确实表示一张曲面, 我们只要假定三个雅可比式

$$\begin{vmatrix}\psi_u & \psi_v\\ \chi_u & \chi_v\end{vmatrix},\quad \begin{vmatrix}\chi_u & \chi_v\\ \varphi_u & \varphi_v\end{vmatrix},\quad \begin{vmatrix}\varphi_u & \varphi_v\\ \psi_u & \psi_v\end{vmatrix} \tag{39b}$$

不同时为零; 用一个公式表示, 我们要求

$$\begin{aligned}(\varphi_u\psi_v-\varphi_v\psi_u)^2&+(\psi_u\chi_v-\psi_v\chi_u)^2\\ &+(\chi_u\varphi_v-\chi_v\varphi_u)^2>0.\end{aligned} \tag{39c}$$

这时, 空间中, 在由 (39a) 所表示的每一点的邻域内, 一定能将三个坐标中的一个用另外两个来表示.

把参变量表示的三个方程 (39a) 用一个向量方程

$$\mathbf{X}=\Phi(u,v) \tag{40a}$$

来代替是方便的, 这里 $\mathbf{X}=(x,y,z)$ 是曲面上点的**位置向量**, Φ 表示向量

$$\Phi(u,v)=(\varphi(u,v),\psi(u,v),\chi(u,v)).$$

曲面上, 在具有参量 u 、 v 的每个点上, 我们能构成**位置向量的偏微商**

$$\mathbf{X}_u=(\varphi_u,\psi_u,\chi_u)\quad 和\quad \mathbf{X}_v=(\varphi_v,\psi_v,\chi_v). \tag{40b}$$

于是向量 $\mathbf{X}$ 的全微分 (见第 51 页公式 (15b)) 是

$$d\mathbf{X}=(dx,dy,dz)=\mathbf{X}_u du+\mathbf{X}_v dv. \tag{40c}$$

---

1) 设 $x=u, y=v$, 我们就看出, 这实际上是参变量形式的特殊情况.

三个行列式 (39b) 恰是向量 $\mathbf{X}_u$ 和 $\mathbf{X}_v$ 的向量积 $\mathbf{X}_u \times \mathbf{X}_v$ 的分量 (见第 196 页). (39c) 左边的表达式表示了向量 $\mathbf{X}_u \times \mathbf{X}_v$ 的长度的平方，所以条件 (39c) 等价于

$$\mathbf{X}_u \times \mathbf{X}_v \neq 0. \tag{40d}$$

例如，半径为 r 的球面 $x^2+y^2+z^2=r^2$, 其参变量表示由方程

$$x = r\cos u \sin v, \quad y = r\sin u \sin v, \quad z = r\cos v \tag{40e}$$

$$(0 \leq u < 2\pi, \quad 0 \leq v \leq \pi)$$

给出，其中 $v=\theta$ 是球面上点的 "余纬", 而 $u=\varphi$ 是球面上点的 "经度" (见 273 页).

这个例子已展示了参变量表示的一种便利. 三个坐标由 u, v 的三个显函数给出，这三个函数都是单值的. 若 v 从 $\frac{\pi}{2}$ 变到 π, 我们得到下半球面，即

$$z = -\sqrt{r^2-x^2-y^2},$$

而当 v 从 0 变到 $\frac{\pi}{2}$ 时得到上半球面. 这样，对参变量表示来说，为了得到整个球面，就不需要像表示法

$$z = \pm\sqrt{r^2-x^2-y^2}$$

那样去考虑这个函数的两个单值分支了.

通过测地投影，我们可得到球面的另一种参变量表示 (见第一卷第 369 页). 为了从北极 $(0,0,r)$ 到赤道平面 $z=0$ 作曲面

$$x^2+y^2+z^2-r^2=0$$

的测地投影，我们把曲面上每一点和北极 N 用直线连接起来，并把这条直线与赤道平面的交点称做曲面上相应点的**测地投影像**(图

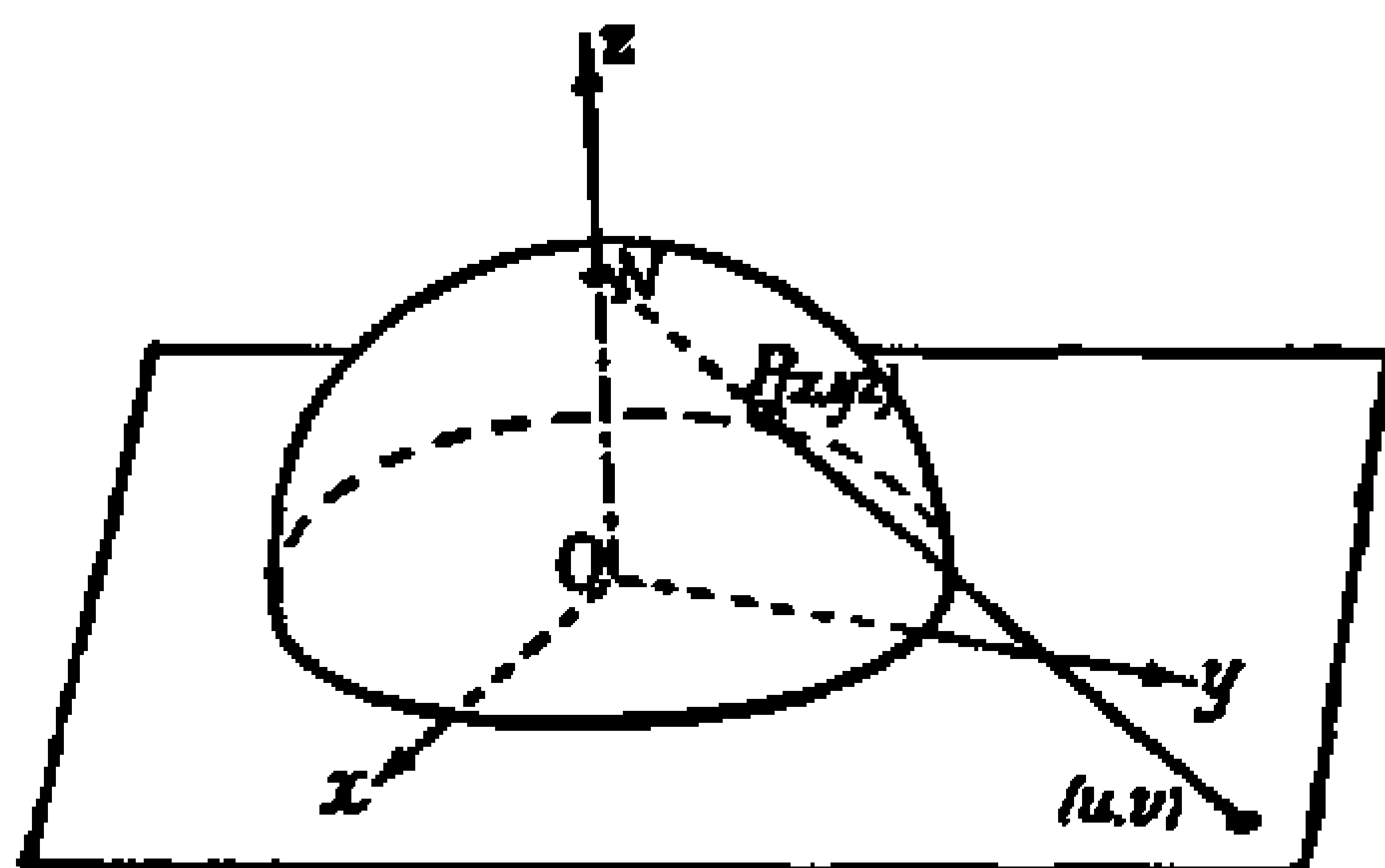


图 3.12 球面的测地投影

3.12). 这样, 除去极点 N 外我们就得到了曲面上的点和平面上的点的一种一一对应. 运用初等几何容易发现, 这个对应由公式

$$x=\frac{2r^2u}{u^2+v^2+r^2},\qquad y=\frac{2r^2v}{u^2+v^2+r^2},$$

$$z=\frac{(u^2+v^2-r^2)r}{u^2+v^2+r^2} \tag{40f}$$

表达, 其中 (u,v) 是平面上像点的直角坐标. 这些方程可看作球面的参变量表示, 参变量 u,v 是 uv 平面上的直角坐标.

作为进一步的例子, 我们给出曲面

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad 和 \quad \frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=-1$$

的参变量表示, 它们分别叫做单叶双曲面和双叶双曲面 (见图 3.13 和 3.14). 单叶双曲面用

$$\begin{aligned} &x=a\cos u\cosh v,\\ &y=a\sin u\cosh v,\\ &z=c\sinh v\\ &(0\le u<2\pi,-\infty<v<+\infty)\end{aligned} \tag{40g}$$

表示; 而双叶双曲面用

$$x=a\cos u\sinh v,$$

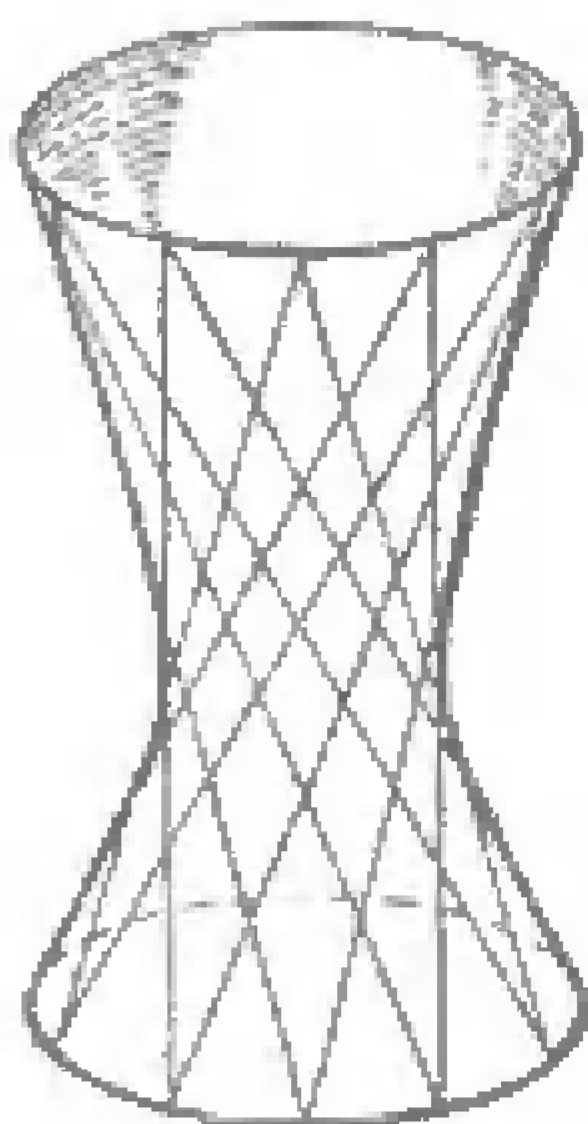


图 3.13 单叶双曲面

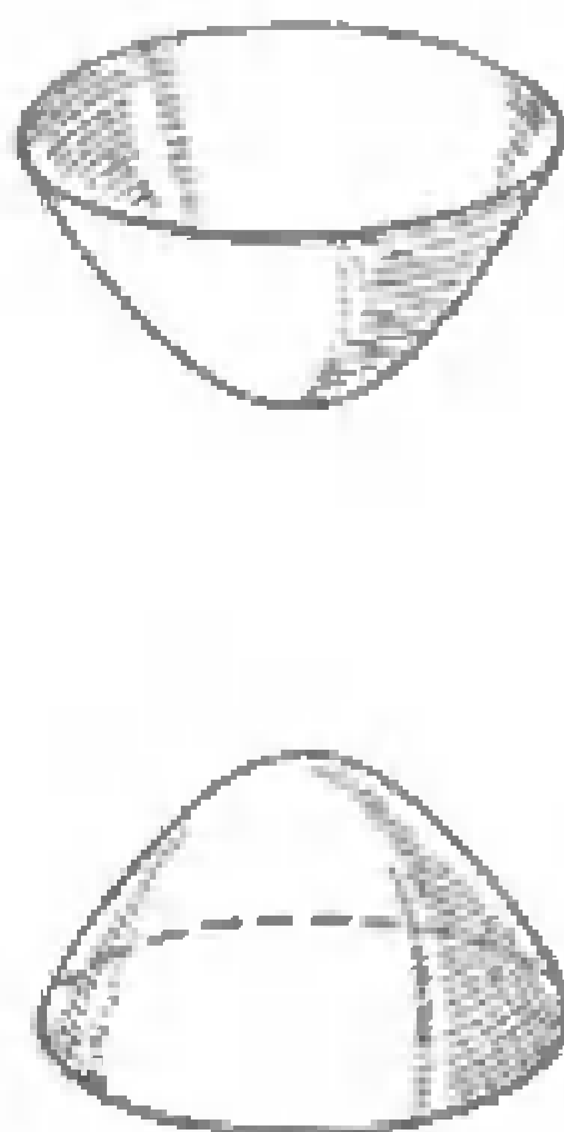


图 3.14 双叶双曲面

$$y = b\sin u \sinh v, \tag{40h}$$

$$z = \pm c\cosh v$$

$$(0 \leq u < 2\pi, 0 < v < +\infty)$$

表示.

一般说来，我们可以把曲面的**参量表示**看作 (u,v) **平面上的区域** R **到相应曲面上的映射**，使 (u,v) 平面上区域 R 的每一点对应于曲面上的一点. 对典型情况反过来也是对的[1].

同样，uv 平面上的曲线 $u = u(t)$, $v = v(t)$ 借助于方程

$$x = \varphi(u(t), v(t)) = x(t), \cdots$$

对应于曲面上的曲线. 例如，在球面的球坐标表示 (40e) 中，方程 $u =$ 常数表示经线，而 $v =$ 常数表示纬线. 如果在我们的参变量表示中，对 u 用一个确定的固定值去代，而让 v 作为参变量，我们就得到一条位于曲面上的 "空间曲线" 或 "挠曲线"; 对 v 用一个固定值去代，而让 u 变也有相应的结论. 所以，在一般情况下我们可以考虑曲面上由方程 $u =$ 常数和 $v=$ 常数所给的那些曲线，这些曲线

1) 当然，不可一概而论. 例如，在球面的球坐标 (见第 306 页) 表示 (40e) 中，球的南北极相应于由 $v = 0$ 和 $v = \pi$ 给出的整个线段.

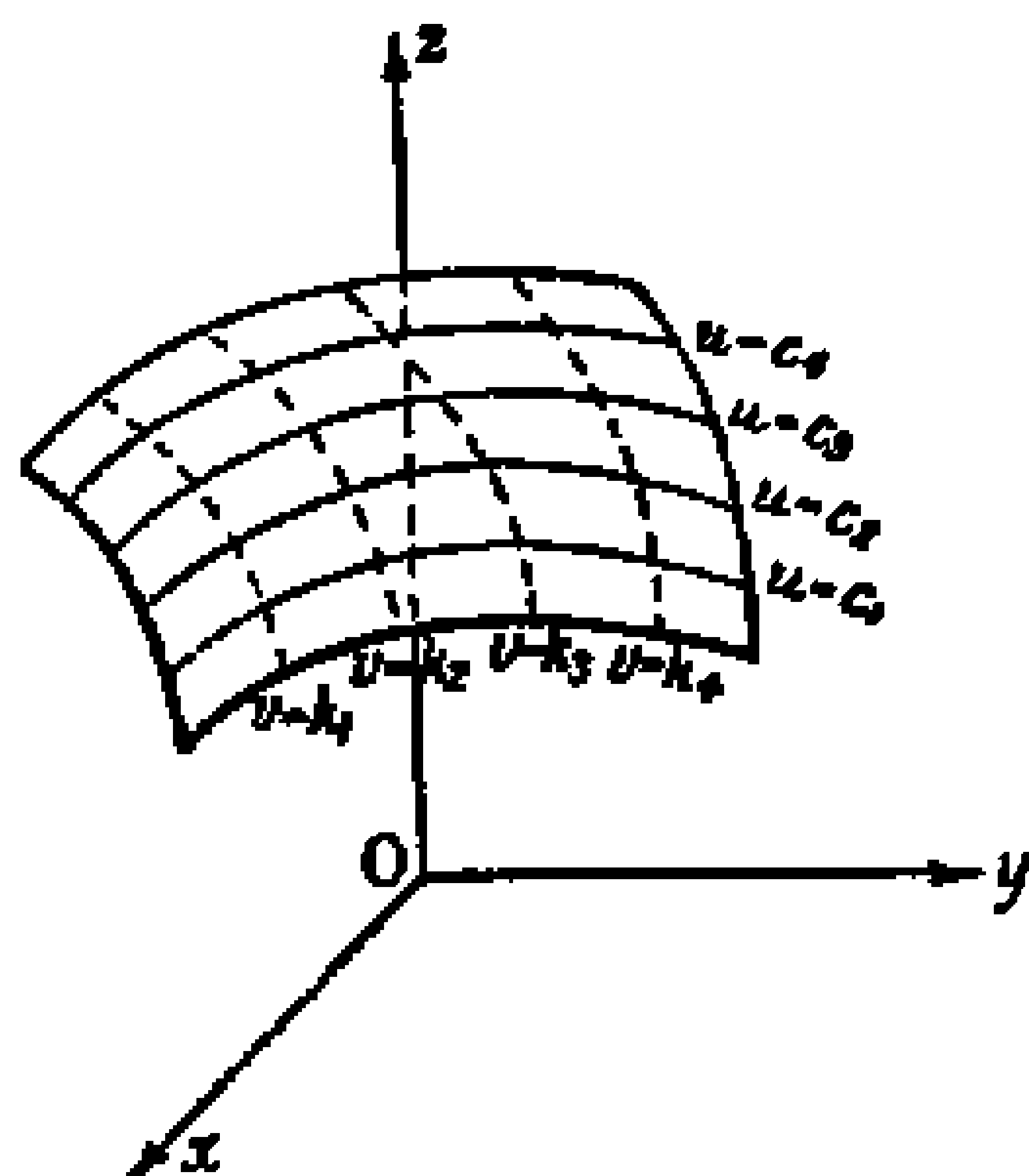


图 3.15 参量曲线 $u=$ 常数, $v=$ 常数

是曲面上的 **参量曲线** 或 **坐标曲线**. 参量曲线网相应于 uv 平面上关于坐标轴的平行线网 (图 3.15).

与 (u,v) 平面上的曲线 $u=u(t)$, $v=v(t)$ 相对应的曲面上的曲线, 其切线的方向与向量

$$
\begin{aligned}
\mathbf{X}_t &= (x_t, y_t, z_t)\\
&= \left(x_u\frac{du}{dt}+x_v\frac{dv}{dt},\quad y_u\frac{du}{dt}+y_v\frac{dv}{dt},\ z_u\frac{du}{dt}+z_v\frac{dv}{dt}\right)\\
&= \mathbf{X}_u\frac{du}{dt}+\mathbf{X}_v\frac{dv}{dt}
\end{aligned}
\tag{41}
$$

的方向相同 (见第 229 页). 在曲面的一个指定点, 曲面上通过该点的一切曲线的切向量 $\mathbf{X}_t$ 依赖于两个向量 $\mathbf{X}_u$, $\mathbf{X}_v$, 它们分别切于通过该点的参量曲线 $u=$ 常数和 $v=$ 常数. 这就意味着, 所有切线都放在通过该点而由向量 $\mathbf{X}_u$ 和 $\mathbf{X}_v$ **张成的** 平面上, 即曲面在该点的 **切平面** 上. 曲面的 **法线** 垂直于所有的切方向, 特别地, 垂直于向量 $\mathbf{X}_u$ 和 $\mathbf{X}_v$. 所以 (见第 197 页) 曲面的法线平行于向量积

$$
\mathbf{X}_u\times\mathbf{X}_v=(y_uz_v-y_vz_u,\ z_ux_v-z_vx_u, x_uy_v-x_vy_u). \tag{42}
$$

探究已给曲面性质的一个最重要的工具是研究曲面上的曲线. 这里我们将只给出这种曲线的弧长 s 的表达式. 在第 229 页 (见第

一卷第 398 页) 曾提到

$$\left(\frac{ds}{dt}\right)^2=\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2=\mathbf{X}_t\cdot\mathbf{X}_t,$$

所以由方程 (41), 我们得到

$$\begin{aligned}\left(\frac{ds}{dt}\right)^2&=\left(\mathbf{X}_u\frac{du}{dt}+\mathbf{X}_v\frac{dv}{dt}\right)\cdot\left(\mathbf{X}_u\frac{du}{dt}+\mathbf{X}_v\frac{dv}{dt}\right)\qquad(43)\\&=\left(x_u\frac{du}{dt}+x_v\frac{dv}{dt}\right)^2+\left(y_u\frac{du}{dt}+y_v\frac{dv}{dt}\right)^2\\&\quad+\left(z_u\frac{du}{dt}+z_v\frac{dv}{dt}\right)^2\\&=E\left(\frac{du}{dt}\right)^2+2F\frac{du}{dt}\frac{dv}{dt}+G\left(\frac{dv}{dt}\right)^2.\end{aligned}$$

这里系数 E,F,G 称为曲面的 **高斯基本量**:

$$E=\left(\frac{\partial x}{\partial u}\right)^2+\left(\frac{\partial y}{\partial u}\right)^2+\left(\frac{\partial z}{\partial u}\right)^2=\mathbf{X}_u\cdot\mathbf{X}_u,\qquad(44a)$$

$$F=\frac{\partial x}{\partial u}\frac{\partial x}{\partial v}+\frac{\partial y}{\partial u}\frac{\partial y}{\partial v}+\frac{\partial z}{\partial u}\frac{\partial z}{\partial v}=\mathbf{X}_u\cdot\mathbf{X}_v,\qquad(44b)$$

$$G=\left(\frac{\partial x}{\partial v}\right)^2+\left(\frac{\partial y}{\partial v}\right)^2+\left(\frac{\partial z}{\partial v}\right)^2=\mathbf{X}_v\cdot\mathbf{X}_v.\qquad(44c)$$

这些系数仅依赖于曲面本身及其参量表示, 而不依赖于曲面上曲线的特殊选择. 表达式 (43) 是弧长 S 关于参量 t 的微商, 在符号上它通常可以写成与沿曲线所用的参量无关, 用二次微分形式 ("基本形式")

$$ds^2=Edu^2+2Fdudv+Gdv^2\qquad(45)$$

给出, ds 称为 **线元素**.

向量积 $\mathbf{X}_u\times\mathbf{X}_v$ 的长度可用 E,F,G 表示, 因为 (见第 197 页)

$$|\mathbf{X}_u\times\mathbf{X}_v|^2=|\mathbf{X}_u|^2|\mathbf{X}_v|^2-(\mathbf{X}_u\cdot\mathbf{X}_v)^2=EG-F^2.\qquad(45a)$$

这样一来，前面用参量表示的原始假定 (39c) 或 (40d) 现在可用关于基本量的条件

$$EG - F^2 > 0 \tag{46}$$

来表示了.

由于单位向量

$$\frac{1}{|\mathbf{X}_u \times \mathbf{X}_v|}\mathbf{X}_u \times \mathbf{X}_v = \frac{1}{\sqrt{EG - F^2}}\mathbf{X}_u \times \mathbf{X}_v$$

的分量是曲面的两个法方向之一的方向余弦，所以，由 (42) 可推出，用参量表示的曲面的法方向有方向余弦

$$\cos\alpha = \frac{y_u z_v - y_v z_u}{\sqrt{EG - F^2}}, \quad \cos\beta = \frac{z_u x_v - z_v x_u}{\sqrt{EG - F^2}},$$

$$\cos\gamma = \frac{x_u y_v - x_v y_u}{\sqrt{EG - F^2}}. \tag{47}$$

曲面上曲线 $u = u(t)$, $v = v(t)$ 的切线与向量

$$\mathbf{X}_t = \mathbf{X}_u \frac{du}{dt} + \mathbf{X}_v \frac{dv}{dt}$$

同方向. 现在我们考虑曲面上以 τ 为参变量数的第二条曲线 $u = u(\tau)$, $v = v(\tau)$. 它的切线与向量

$$\mathbf{X}_\tau = \mathbf{X}_u \frac{du}{d\tau} + \mathbf{X}_v \frac{dv}{d\tau}$$

同方向. 若这两条曲线通过曲面上同一点，则其交角 ω 的余弦与向量 $\mathbf{X}_t$ 与 $\mathbf{X}_\tau$ 之间夹角的余弦一样. 因此 (见第 141 页)

$$\cos\omega = \frac{\mathbf{X}_t \cdot \mathbf{X}_\tau}{|\mathbf{X}_t|\,|\mathbf{X}_\tau|}.$$

在这里

$$\mathbf{X}_t \cdot \mathbf{X}_\tau = \left(\mathbf{X}_u \frac{du}{dt} + \mathbf{X}_v \frac{dv}{dt}\right) \cdot \left(\mathbf{X}_u \frac{du}{d\tau} + \mathbf{X}_v \frac{dv}{dt}\right)$$

$$= E\frac{du}{dt}\frac{du}{d\tau} + F\left(\frac{du}{dt}\frac{dv}{d\tau} + \frac{du}{d\tau}\frac{dv}{dt}\right) + G\frac{dv}{dt}\frac{dv}{d\tau}.$$

所以曲面上两曲线间夹角的余弦由下式给出:

$$\cos\omega = \left[E\frac{du}{dt}\frac{du}{d\tau} + F\left(\frac{du}{dt}\frac{dv}{d\tau} + \frac{du}{d\tau}\frac{dv}{dt}\right) + G\frac{dv}{dt}\frac{dv}{d\tau}\right] \Big/ \left[\sqrt{E\left(\frac{du}{dt}\right)^2 + 2F\frac{du}{dt}\frac{dv}{dt} + G\left(\frac{dv}{dt}\right)^2} \times \sqrt{E\left(\frac{du}{d\tau}\right)^2 + 2F\frac{du}{d\tau}\frac{dv}{d\tau} + G\left(\frac{dv}{d\tau}\right)^2}\right]. \tag{48}$$

一个平面区域到另一个平面区域上的映射可以看作参量表示的一种特殊情况, 因为如果假定 (39a) 中的第三个函数 $\chi(u,v)$ 对所考虑的 (u,v) 恒为零, 则所述方程就只表示了 uv 平面上的区域到 xy 平面上的区域的映射. 如果我们愿意用坐标变换的术语来说的话就是, 方程组确定了 uv 区域内的一个曲线坐标系, 而反函数 (要是它们存在的话) 则确定了 xy 平面内的一个 uv 曲线坐标. 利用这种坐标 (u,v) 来表示, xy 平面上的线元素可简单地写为 [见 (44a, b, c)]

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2,$$

其中

$$E = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2, \tag{49a}$$

$$F = \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v}, \tag{49b}$$

$$G = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2. \tag{49c}$$

我们取**环面**作为曲面的参量表示的一个更深入的例子. 环面是这样得到的: 取一个圆, 让它绕着一条与它在同一个平面上而又与它不相交的直线旋转 (见图 3.16). 我们取旋转轴作为 z 轴, y 轴

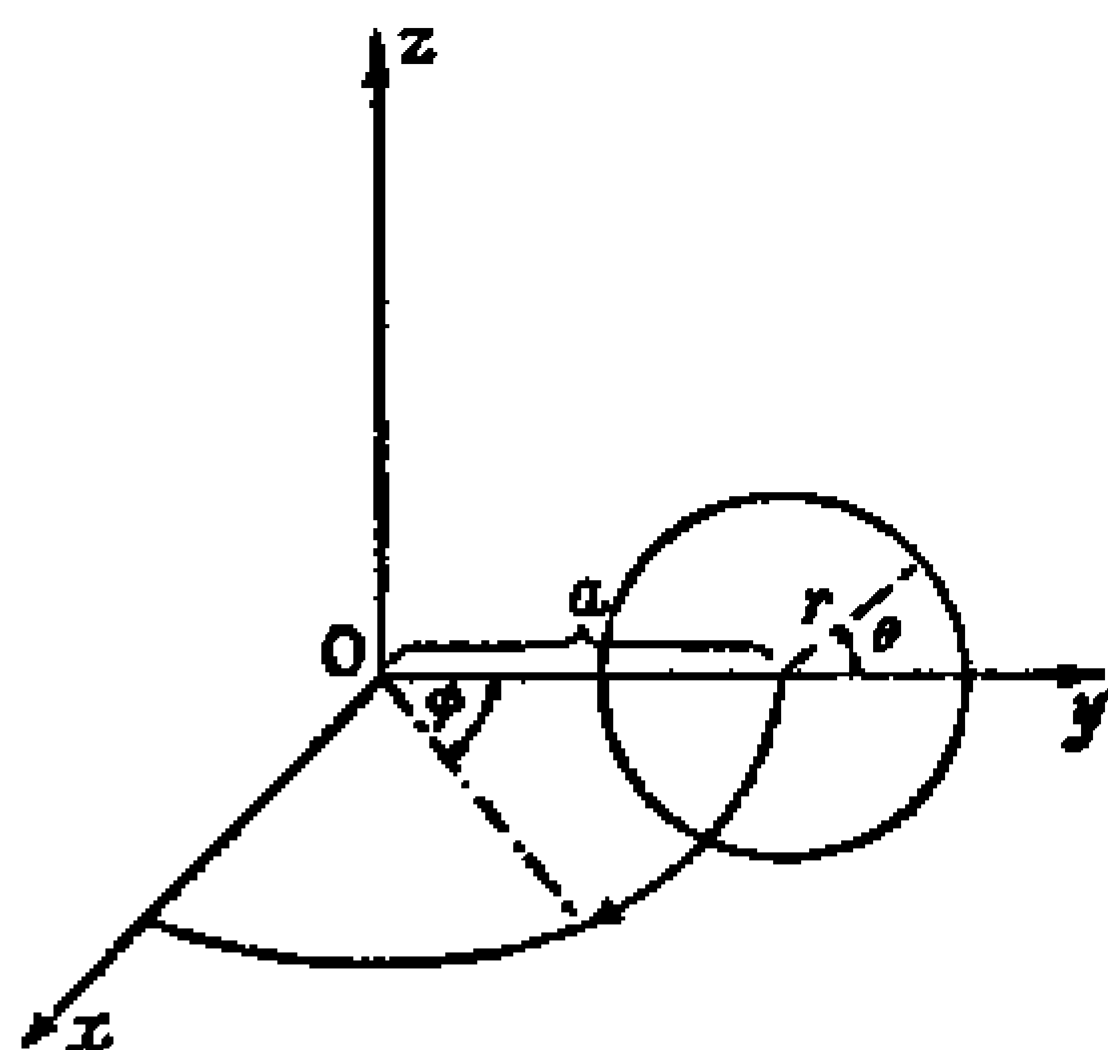


图 3.16 旋转圆产生环面

通过该圆的中心，中心的 y 坐标为 a. 如果圆的半径是 $r < |a|$, 我们就得到该圆在 yz 平面上的参量表示:

$$x = 0, \quad y - a = r\cos\theta, \quad z = r\sin\theta \quad (0 \le \theta < 2\pi).$$

现在让圆绕 z 轴旋轴，我们发现对圆上每一个点而言 $x^2 + y^2$ 保持常值，也就是 $x^2 + y^2 = (a + r\cos\theta)^2$. 如果 φ 是关于 z 轴的转角，则

$$x = (a + r\cos\theta)\sin\varphi,$$

$$y = (a + r\cos\theta)\cos\varphi,$$

$$z = r\sin\theta$$

$$(0 \le \varphi < 2\pi, \quad 0 \le \theta < 2\pi).$$

这就是环面的以 θ 和 φ 为参量的参量表示. 在这种表示中，环面成了 $\theta\varphi$ 平面上的边长为 2π 的正方形的像，其中位于同一直线 $\theta=$ 常数上或 $\varphi =$ 常数上的每对边界点仅对应于曲面上的一个点，而正方形的四个顶点只对应于同一个点.

对于环面的线元素，依 (44a,b,c), 我们有

$$ds^2 = r^2 d\theta^2 + (a + r\cos\theta)^2 d\varphi^2.$$

练　习　3.4 a

1. 计算下列曲面上的线元素:

(a) 球面

$$x=\cos u\sin v,\quad y=\sin u\sin v,\quad z=\cos v;$$

(b) 双曲面

$$x=\cos u\cosh v,\quad y=\sin u\cosh v,\quad z=\sinh v;$$

(c) 由

$$r=\sqrt{x^2+y^2}=f(z)$$

所给出的旋转曲面 (用柱坐标 z 和 $\theta=\operatorname{arctg}(y/x)$ 作为曲面上的坐标);

(d) 由

$$\frac{x^2}{a-t}+\frac{y^2}{b-t}+\frac{z^2}{c-t}=1$$

所给出的共焦二次曲面族中相当于 t_3= 常数的二次曲面, 取 t_1,t_2 作为二次曲面的坐标 (见第 280 页习题 9).

2. 对悬链曲面

$$x=a\cosh\left(\frac{t}{a}\right)\cos\left(\frac{\theta}{a}\right),\quad y=a\cosh\left(\frac{t}{a}\right)\sin\left(\frac{\theta}{a}\right),$$

$z=t$, 求其高斯基本量; 证明 $E-G=F=0$.

3. 对曲面 $x=u\cos v, y=u\sin v, z=\alpha u+\beta$ ($\alpha,\beta=$ 常数) 证明: 曲线 $u=$ 常数, $v=$ 常数的象是正交的.

4. 对于方程 $z=f(x,y)$ 所给出的曲面, 确定其线元素的基本型是什么?

5. 如果在以 u,v 为参量的曲面上, 通过方程 $u=u(s,t)$, $v=v(s,t)$ 引进新的曲线坐标系 (s,t), 试证明

$$E'G'-F'^2=(EG-F^2)\left\{\frac{d(u,v)}{d(r,s)}\right\}^2,$$

其中 E', F', G' 是关于 s, t 所取的基本量，而 E, F, G 是关于 u, v 所取的基本量.

6. 设 t 是曲面 S 在点 P 的切线，考虑一切包含 t 的平面与曲面的交线. 证明所有这些不同交线的曲率中心都落在同一个圆上.

7. 若 t 是曲面 S 在点 P 的一条切线，则我们把在 P 点通过 t 的法平面（即通过 t 和法线的平面）与曲面的交线的曲率称为 **S 在方向 t 的曲率** κ. 对于在 P 点的每一切线，我们做一个以 P 为始点，以 t 为方向的向量，其长度为 $1/\sqrt{\kappa}$. 证明所有这些向量的终点都落在同一个二次曲线上.

8. 设一曲线是作为两个曲面

$$x^2 + y^2 + z^2 = 1$$

$$ax^2 + by^2 + cz^2 = 0$$

的交线给出的，试求

(a) 切线方程，

(b) 在曲线上任一点处密切平面的方程.

9. 若球面上点的坐标 (x, y, z) 由方程 (见第 272 页)

$$x = a\sin\theta\cos\varphi, \quad y = a\sin\theta\sin\varphi, \quad z = a\cos\theta$$

给出，试证明，通过任何一点 (θ, φ) 的两族曲线 $\theta+\varphi = \alpha, \theta-\varphi = \beta$ 之间的交角为 $\arccos\{(1-\sin^2\theta)/(1+\sin^2\theta)\}$ (见第 311 页).

证明，任一条曲线的曲率半径等于

$$\frac{a(1+\sin^2\theta)^{\frac{3}{2}}}{(5+3\sin^2\theta)^{\frac{1}{2}}}.$$

b. 一般保角变换

如果平面上的一个变换

$$x = \varphi(u, v), \quad y = \psi(u, v) \tag{50}$$

把任何两条相交的曲线变为两条相交的曲线，并保持其夹角不变，则这个变换称为一个保角变换.

定理 一个连续可微的变换(50)是保角的一个必要充分条件是它满足柯西-黎曼方程

$$\varphi_u - \psi_v = 0, \quad \varphi_v + \psi_u = 0 \tag{51a}$$

或

$$\varphi_u + \psi_v = 0, \quad \varphi_v - \psi_u = 0. \tag{51b}$$

在第一种情况，保持角度方向;在第二种情况，角度反向 [1].

证明如下：假定变换是保角的，则在 uv 平面上两族正交的曲线 $u=$ 常数 $=u_0, v=v_0+t$ 与 $u=u_0+\tau, v=$ 常数 $=v_0$ 一定变为 xy 平面上两族正交的曲线. 由关于两曲线交角的公式 (48)(见第 312 页) 直接推出

$$0 = F = \varphi_u\varphi_v + \psi_u\psi_v. \tag{51c}$$

同样，对应于 $u=u_0+t, v=v_0+t$ 和 $u=u_0+\tau, v=v_0-\tau$ 的曲线也一定是正交的. 这就给出

$$0 = E - G = \varphi_u^2 + \psi_u^2 - \varphi_v^2 - \psi_v^2. \tag{51d}$$

方程 (51c) 可写为

$$\varphi_u = \lambda\psi_v, \quad \varphi_v = -\lambda\psi_u,$$

其中 λ 是一个适当的常数. 将此式代入 (51d), 我们立刻得到 $\lambda^2=1$, 所以上述的柯西-黎曼方程组 (51a,b) 中总有一个成立.

从下面的观察可确信，柯西-黎曼方程组是保角性的一个充分条件. 除非四个量 $\varphi_u, \varphi_v, \psi_u, \psi_v$ 全是零.

由方程 (51a) 或 (51b) 可推出关系式

$$E = G \geq 0, \quad F = 0.$$

---

1) 最后的断言直接由第 284 页上关于雅可比式 $D=\varphi_u\psi_v-\varphi_v\psi_u$ 的符号推出. 在 (51a) 的情况下，我们有 $D=\varphi_u^2+\varphi_v^2\geq 0$, 在 (51b) 的情况下 $D=-\varphi_u^2-\varphi_v^2\leq 0$.

E, F, G 是由 (49a,b,c) 所定义的基本量. 于是, 根据 (48), 在 xy 平面上两曲线间的夹角 ω 由下式给出

$$\cos\omega = \frac{\dfrac{du}{dt}\dfrac{du}{d\tau} + \dfrac{dv}{dt}\dfrac{dv}{d\tau}}{\sqrt{\left(\dfrac{du}{dt}\right)^2 + \left(\dfrac{dv}{dt}\right)^2}\sqrt{\left(\dfrac{du}{d\tau}\right)^2 + \left(\dfrac{dv}{d\tau}\right)^2}}.$$

这个方程的右边恰是 uv 平面上相应曲线间夹角的余弦. 所以, 映射保持曲线间夹角的大小, 但可能改变其定向. 只有使 $E = F = G = 0$ 的点例外, 也就是使两个映射函数的所有一阶微商都为零 [1] 的点例外.

练 习 3.4 b

1. 研究映射 $x = u^2 - v^2, y = 2uv$ 的性状. 在 $u = 2, v = 3$ 它是否保角的? 在 $u = v = 0$ 呢? 为什么?

2. 映射 $x = \dfrac{1}{2}\log(u^2 + v^2), y = \arctan\dfrac{v}{u}$ 在什么地方是保角的?

3. 证明, 如果映射 $(u, v) \to (x, y)$ 和 $(u, v) \to (\xi, \eta)$ 都是保角的, 则映射 $(u, v) \to (x\xi - y\eta, x\eta + y\xi)$ 也是保角的.

4. (a) 证明单位球面到平面的测地投影是保角的.

(b) 证明球面上的圆变为平面上的圆或直线.

(c) 证明, 在测地投影中, 球面关于赤道平面的反射对应于 uv 平面上的反演.

(d) 以 u, v 为参变量找出球面上线元素的表达式.

5. 高斯基本系数 (44) 满足什么条件时, 从 uv 平面到曲面 $\mathbf{X} = \mathbf{X}(u, v)$ 的映射才是保角的?

6. 寻找一个从球面 $x = \cos\theta\sin\varphi, y = \sin\theta\sin\varphi, z = \cos\varphi$ 到 uv 平面的保角变换, 使得 $\theta = u, \varphi = f(v)$, 其中 $f(v)$ 满足 $f(0) = \dfrac{1}{2}$.

---

1) 在这种地方映射实际上不再是保角的了.

3.5 曲线族，曲面族，以及它们的包络

a. 一般说明

在许多场合中，我们已经不把曲线或曲面作为一个单独的个体来考虑，而是作为一族曲线或一族曲面中的一个成员来考虑，如像 $f(x,y)=c$，这里对每一个 c 都存在着族中一条曲线与它对应.

例如，在 xy 平面上与 y 轴平行的直线，即 $x=c$ 形成一个曲线族. 关于原点的同心圆族 $x^2+y^2=c^2$ 也一样；对于 c 的每一个值都对应着族中一个圆，即半径为 c 的那个圆. 类似地，等轴双曲线 $xy=c$ 形成一个曲线族，其略图如图 3.2. 特殊值 $c=0$ 对应于由两条坐标轴构成的退化双曲线. 给定曲线的一切法线所成的集合是曲线族的另一个例子. 若曲线以 t 为参变量而由方程 $\xi=\varphi(t)$，$\eta=\psi(t)$ 给出，则我们可得到形如（见第一卷第 388 页）

$$(x-\varphi(t))\varphi'(t)+(y-\psi(t))\psi'(t)=0$$

的法线族方程，这里 t 代替 c 而表示曲线族的参变量.

曲线族的一般概念可用解析方法表述如下. 设

$$f(x,y,c)$$

是一个含有两个自变量 x 和 y 及一个参变量 c 的连续可微函数，其中参变量 c 在给定区间内变化.（这样，参变量实际上是第三个自变量，仅仅因为它起着不同的作用，才把字母写得有区别.）如果对参变量 c 的每一个值，方程

$$f(x,y,c)=0 \tag{52a}$$

表示一条曲线，则 c 在它的区间上变化时所得曲线的全体叫做依赖于参变量 c 的**曲线族**.

族中每一条曲线也可表为参变量形式：

$$x=\varphi(t,c),\quad y=\psi(t,c), \tag{52b}$$

其中 c 是区分族中不同曲线的参变量，而 t 是沿曲线的参变量.

例如，方程组

$$x = c\cos t, \quad y = c\sin t$$

表示上面提到的同心圆族；又如方程组

$$x = ct, \qquad y = \frac{1}{t},$$

表示上面提到的正交双曲线族，但除去由两条坐标轴构成的退化双曲线.

偶而我们也被引导到去考虑依赖于几个参变量的曲线族. 例如，在平面上一切圆 $(x-a)^2+(y-b)^2=c^2$ 的全体是依赖于三个参变量 a, b, c 的曲线族. 如果在叙述中不出现相反情况的话，我们总把曲线族理解为依赖于一个参变量的"单参变量"的曲线族. 其他情况我们将有区分地说两参变量，三参变量，多参变量的曲线族.

关于空间中的曲面族，类似的论述当然也成立. 如果给了一个连续可微函数 $f(x, y, z, c)$，并且如果对某个确定区间上的参变量 c 的每一个值，方程

$$f(x, y, z, c) = 0$$

表示空间中的一张以 (x, y, z) 为直角坐标的曲面，则令 c 在它的区间内变化时所得曲面的全体称为一个**曲面族**，或更确切些说，是一个以 c 为**参变量的单参变量曲面族**. 例如，中心在原点的球面 $x^2+y^2+z^2=c^2$ 就形成这样一个曲面族. 和曲线族的情况一样，我们也考虑依赖于几个参变量的曲面族.

例如，由方程

$$ax + by + \sqrt{1-a^2-b^2}\,z + 1 = 0$$

所确定的诸平面形成依赖于参变量 a 和 b 的两参变量曲面族，参变量 a 和 b 在区域 $a^2+b^2 \le 1$ 内变化. 这个曲面族是由与原点距离为 1 的所有平面构成的 [1].

---

1) 有时一个单参变量的曲面族被归入 ∞^1 曲面，一个双参变量曲面族被归入 ∞^2 曲面，等等.

练 习 3.5 a

1. 从几何上刻画下列的曲线族:

(a) $\frac{x^2}{a^2}+\frac{y^2}{b^2}=c^2$, a,b 为已知常数, c 是参变量;

(b) $x^2+(y-c)^2=c^2$, c 是参变量;

(c) $x=\cos(c+t)$, $y=\sin(c+t)$, $0\leq t\leq 2\pi$, c 是参变量.

2. 描述单参变量曲面族

$$(x-c)^2+(y-1-c)^2+(z+\sqrt{2}-2c)^2=1.$$

b. 单参量曲线的包络

如果某一直线族由一平面曲线 E 的诸切线构成 (例如, 一条曲线 C 的法线族就是 C 的包络 E 的切线族, 见第一卷), 这时我们就说**曲线 E 是这一直线族的包络**. 同样, 我们将说半径为 1, 中心在 x 轴上的圆族 —— 即具有方程 $(x-c)^2+y^2-1=0$ 的圆族 —— 以直线对 $y=1$ 和 $y=-1$ 为其包络, 该直线与族中每一个圆相切 (图 3.17). 在这两例中, 找出族中以 c 和 $c+h$ 为参数的两条曲线的交点, 而后让 h 趋近于 0, 我们就能得到包络和族中以 c 为参数的那条曲线的交点. 我们简短地用下述语言来表示: 包络是**相邻曲线交点的轨迹**.

对任何一个曲线族, 如果曲线 E **在其每一点与族中某一曲线相切, 则称它为曲线族的包络**. 现在产生了寻找已给曲线族 $f(x, y, c)=0$ 的包络 E 的问题. 我们先作些表面上说得通的说明, 在说明中, 我们假定包络确实存在, 并像上面的情形那样, 它能作为相邻曲线的交点的轨迹而得到. [1] 于是我们就通过下述途径来得

1) 因为有例子表明后面这一假定限制太强了, 我们即将用一个更完全的讨论来代替这个似真性的讨论.

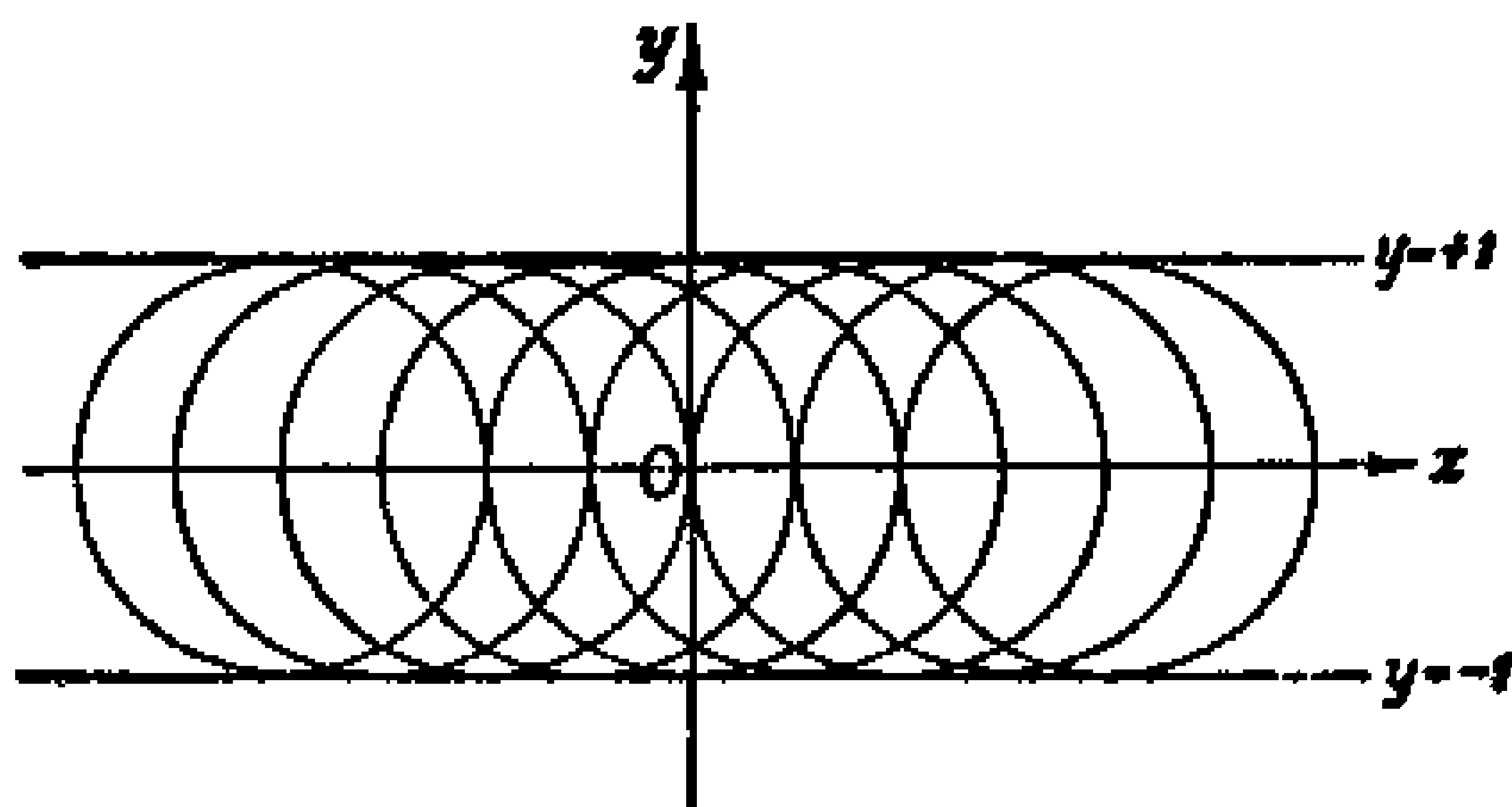


图 3.17　具有包络的圆族

到曲线 $f(x,y,c)=0$ 和曲线 E 的接触点：除了这条曲线外，我们还考虑一条相邻曲线 $f(x,y,c+h)=0$，找出这两条曲线的交点，然后让 h 趋近于 0. 这时交点一定趋近于所要找的点. 在交点处除了方程

$$\frac{f(x,y,c+h)-f(x,y,c)}{h}=0$$

成立外，还有方程 $f(x,y,c+h)=0$ 与方程 $f(x,y,c)=0$ 也成立. 在第一个方程中，让 $h\to 0$ 取极限. 由于我们假定了偏微商 f_c 的存在性，这就给出了关于曲线 $f(x,y,c)=0$ 与包络的接触点的两个方程

$$f(x,y,c)=0,\quad f_c(x,y,c)=0. \tag{53}$$

如果我们根据这些方程把 x 和 y 确定为 c 的函数，我们就得到一条以 c 为参变量的曲线的参量表达式，而这条曲线就是包络. 消去参变量 c，曲线也能表成 $g(x,y)=0$ 的形式. 这个方程叫做曲线族的**判别式**，由方程 $g(x,y)=0$ 给定的曲线叫做**判别曲线**.

于是我们得到下述法则：**为了得到曲线族** $f(x,y,c)=0$ **的包络，我们考虑联立方程组** $f(x,y,c)=0$ **和** $f_c(x,y,c)=0$，**并试图通过方程组将** x **和** y **表为** c **的函数，或在这两方程间消去参变量** c.

现在我们依据包络作为接触点的曲线的定义来作更一般的讨论，而代替上面的直观考虑. 同时，我们要弄清楚在什么条件下我

们的法则在实际上给出包络，并且有什么其他可能会出现.

首先，我们假定 E 是包络，它能利用参变量 c 通过两个连续可微函数

$$x = x(c), \qquad y = y(c)$$

表示出来，在这里

$$\left(\frac{dx}{dc}\right)^2 + \left(\frac{dy}{dc}\right)^2 \neq 0;$$

我们还假定 E 在具有参变量 c 的点与曲线族 $f(x,y,c) = 0$ 中有同参变量值 c 的曲线相切. 于是接触点满足方程 $f(x,y,c) = 0$. 因此，如果我们在这个方程中用 $x(c)$ 和 $y(c)$ 代替 x 和 y, 则对区间中的一切 c 值方程仍能成立. 关于 c 求微商，我们立刻得到

$$f_x \frac{dx}{dc} + f_y \frac{dy}{dc} + f_c = 0.$$

现在相切的条件是

$$f_x \frac{dx}{dc} + f_y \frac{dy}{dc} = 0,$$

因为量 $\frac{dx}{dc}$ 和 $\frac{dy}{dc}$ 与 E 的切线的方向余弦成比例，且量 f_x 和 f_y 与族中曲线 $f(x,y,c) = 0$ 的法线的方向余弦成比例，而这些方向又一定彼此交成直角. 这就推出包络满足方程 $f_c = 0$, 于是我们看到了，方程 (53) 构成了包络的一个 **必要条件**.

为了找出这个条件什么时候还是 **充分** 的，我们假定由两个连续可微函数 $x = x(c)$ 和 $y = y(c)$ 所表示的曲线 E 满足两个方程 $f(x,y,c) = 0$ 和 $f_c(x,y,c) = 0$. 在 $f(x,y,c) = 0$ 中我们再用 $x(c)$ 和 $y(c)$ 代 x 和 y; 这时这个方程变成了 c 的恒等式. 如果我们关于 c 取微商，并记着 $f_c = 0$, 我们立刻得到关系式

$$f_x \frac{dx}{dc} + f_y \frac{dy}{dc} = 0,$$

这个关系式对所有的 c 成立. 如果两个表达式

$$f_x^2 + f_y^2 \quad \text{和} \quad \left(\frac{dx}{dc}\right)^2 + \left(\frac{dy}{dc}\right)^2$$

都在 E 的一点异于 0, 因而在这点曲线 E 和族中曲线都有确定的切线, 则这个方程说明包络与族中曲线彼此相切. 加上这些补充假定, 我们的法则就不仅是包络的一个必要条件而且也是包络的一个充分条件. 但是, 如果 f_x 和 f_y 都为 0, 则族中的曲线可能有一个奇点 (见第 256 页), 而关于曲线的相切我们引不出任何结论.

于是, 在我们找到了判别曲线之后, 还必须进一步研究每一种情况, 以便确定它是否确实是一个包络或者在什么范围内它不是.

最后, 我们叙述以 t 为参变量的、用参变方程

$$x=\varphi(t,c),\qquad y=\psi(t,c)$$

形式给定的曲线族的判别曲线的条件. 这条件是

$$\varphi_t\psi_c-\varphi_c\psi_t=0.$$

消去 t, 从曲线族的参变量表示过渡到原始表示, 我们就能够立即得出这一条件.

练 习 3.5 b

1. 光滑平面曲线的法线族总有包络吗?

2. 直线族

$$y=cx+\psi(c)$$

满足微分方程

$$y=xy'+\psi(y')$$

(克莱罗方程). 求该直线族的包络的非参量方程, 并验证它也一定满足此微分方程.

c. 例

1. $(x-c)^2+y^2=1$. 正如我们在第 320 页上看到的, 这个方程表示中心在 x 轴上半径为 1 的圆族 (图 3.17). 从几何上我们立刻

看出，包络一定是由两条直线 $y=1$ 和 $y=-1$ 构成的. 我们可依据我们的法则来验证，因为方程 $(x-c)^2+y^2=1$ 和 $-2(x-c)=0$ 立刻给我们以形如 $y^2=1$ 的包络.

2. 半径为 1 通过原点的圆族，其中心一定位于以 1 为半径以原点为中心的圆上，并由方程

$$(x-\cos c)^2+(y-\sin c)^2=1,$$

或

$$x^2+y^2-2x\cos c-2y\sin c=0$$

给出. 关于 c 取微商即得 $x\sin c-y\cos c=0$. 这两个方程为 $x=0$ 和 $y=0$ 所满足. 但是，如果 $x^2+y^2\neq 0$, 从我们的方程容易推出 $\sin c=\frac{y}{2}$, $\cos c=\frac{x}{2}$, 从而消去 c 我们就得到 $x^2+y^2=4$. 于是，正像我们从几何直观上所能预期的那样，我们的法则给出了半径为 2 以原点为心的圆作为包络，但是它也给出了我们一个孤立点 $x=0$, $y=0$.

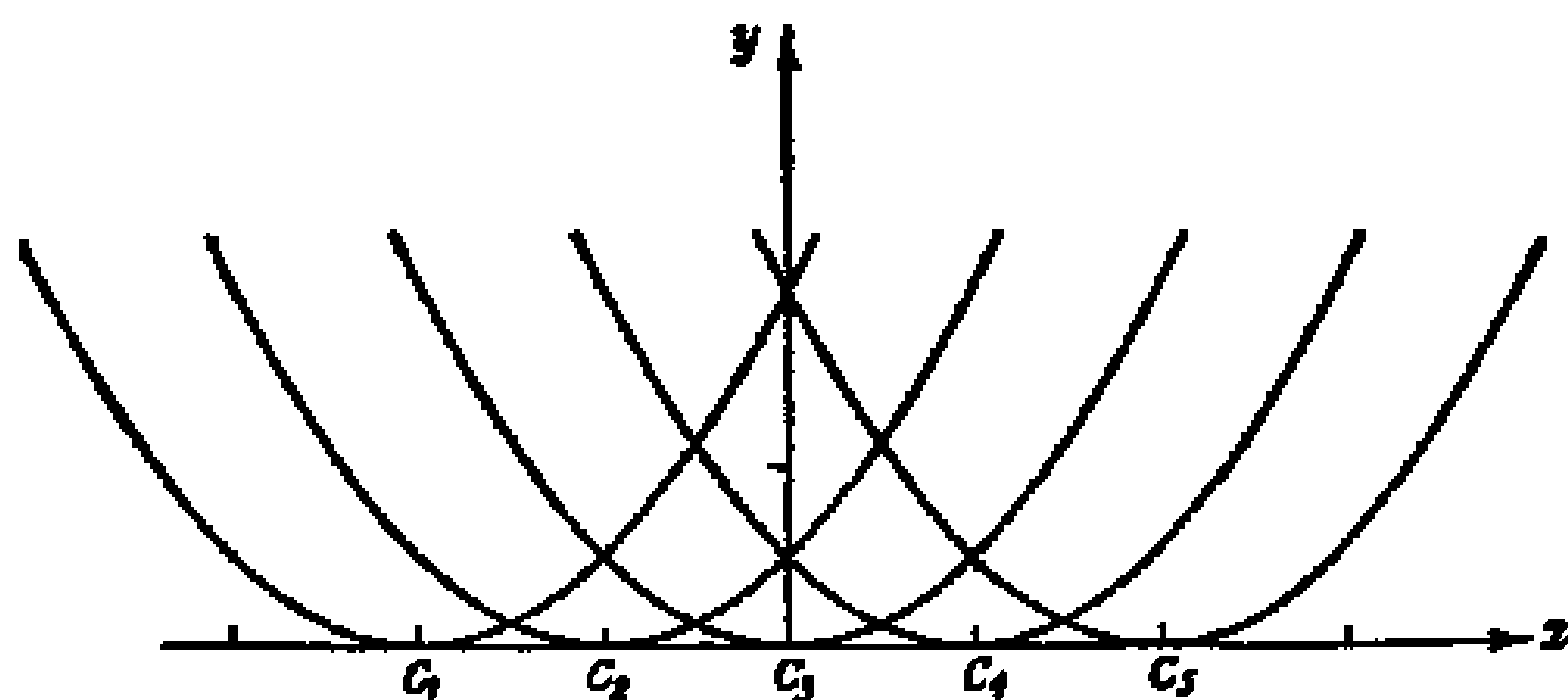


图 3.18　具有包络的抛物线族

3. 抛物线族 $(x-c)^2-2y=0$(见图 3.18) 也有一个包络，这用几何直观和用我们的法则都能找出，就是 x 轴.

4. 我们考虑圆族 $(x-2c)^2+y^2-c^2=0$(见图 3.19). 关于 c 求微商得到 $2x-3c=0$, 代入原方程我们找到包络的方程为

$$y^2=\frac{x^2}{3};$$

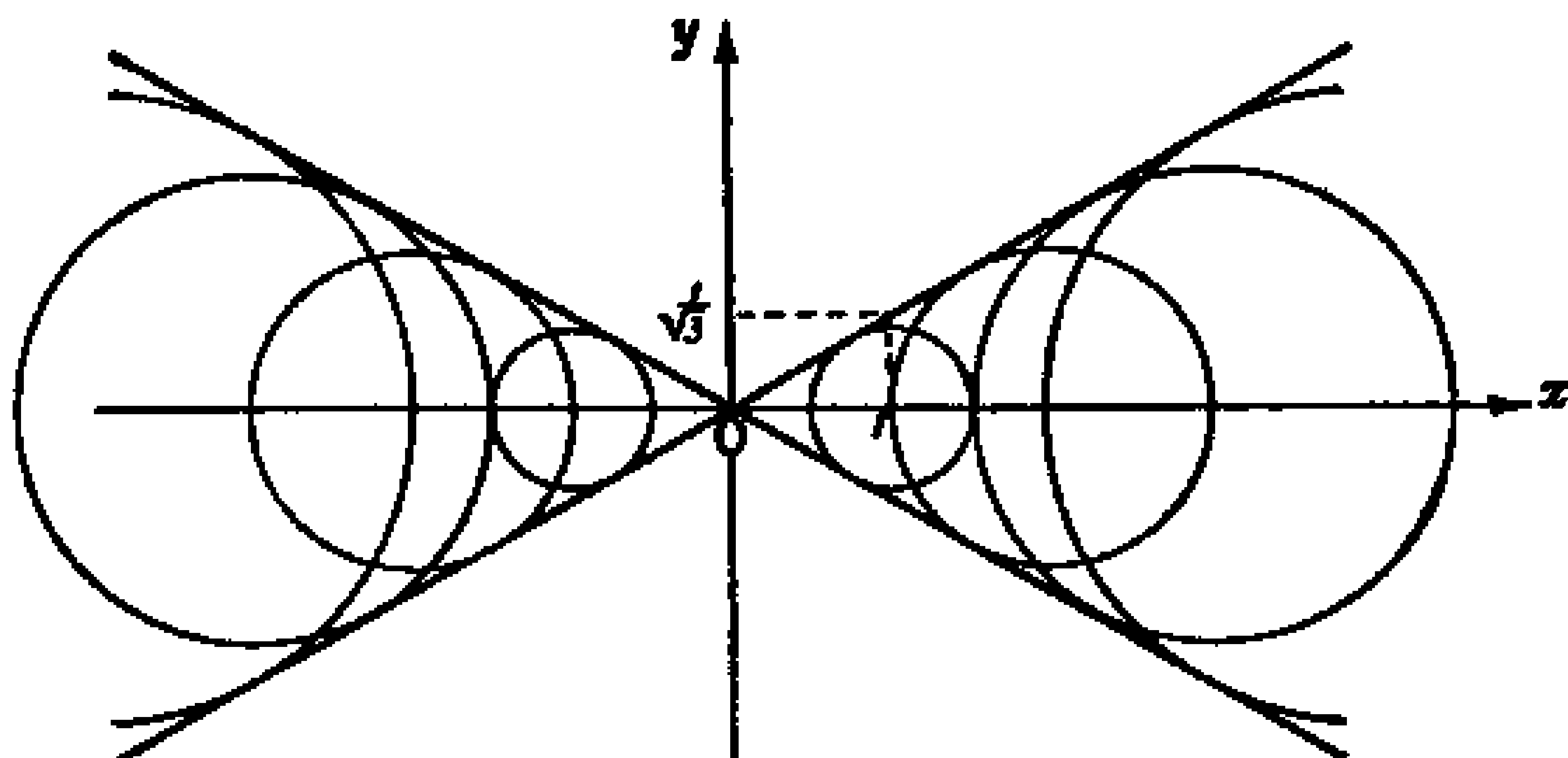


图 3.19　圆族 $(x-2c)^2+y^2-c^2=0$

这是说，包络由两条直线

$$y=\frac{1}{\sqrt{3}}x \quad 和 \quad y=-\frac{1}{\sqrt{3}}x$$

构成. 原点是个例外值，它不是切点.

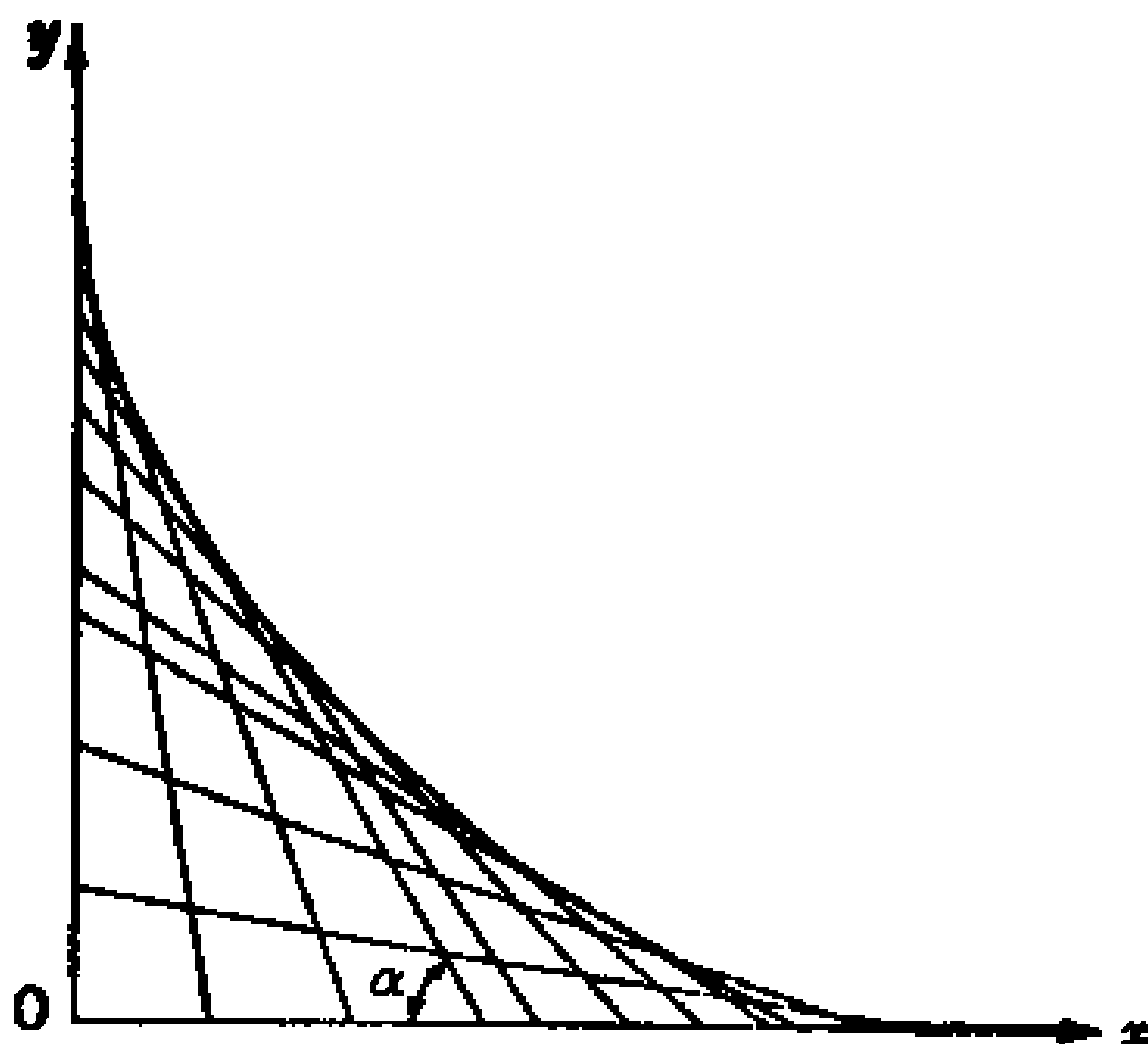


图 3.20　星形线的一段弧作为直线族的包络

5. 我们下面考虑由 x 轴和 y 轴切出单位长度的直线族. 如果 $\alpha=c$ 是图 3.20 中标出的角，则直线族由方程

$$\frac{x}{\cos\alpha}+\frac{y}{\sin\alpha}=1$$

给出. 包络的条件是

$$\frac{\sin\alpha}{\cos^2\alpha}x-\frac{\cos\alpha}{\sin^2\alpha}y=0,$$

它与直线族的方程一起给出参变量形式的包络

$$x=\cos^3\alpha,\qquad y=\sin^3\alpha.$$

消去参变量, 我们得到方程

$$x^{2/3}+y^{2/3}=1.$$

这条曲线叫做星形线 (见第一卷第四章习题 1, 第 485 页). 它由交于四个尖点的四条对称的分枝构成 (图 3.21 和 3.22).

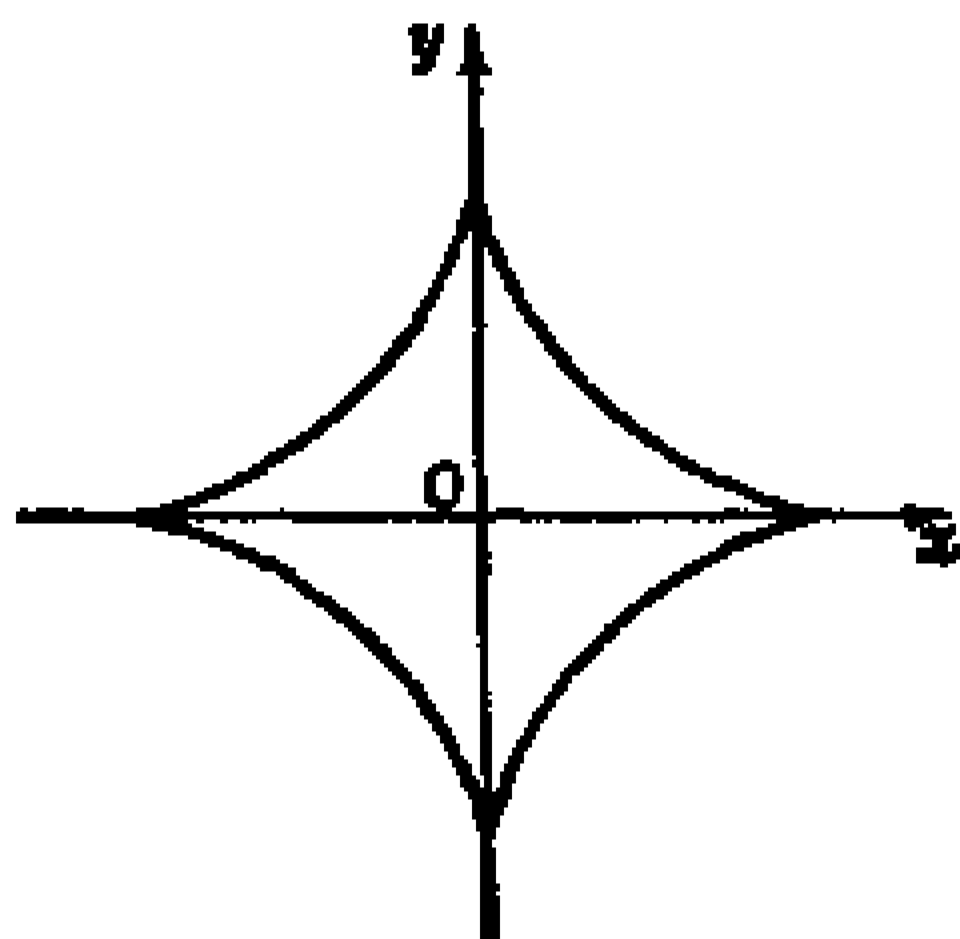


图 3.21 星形线

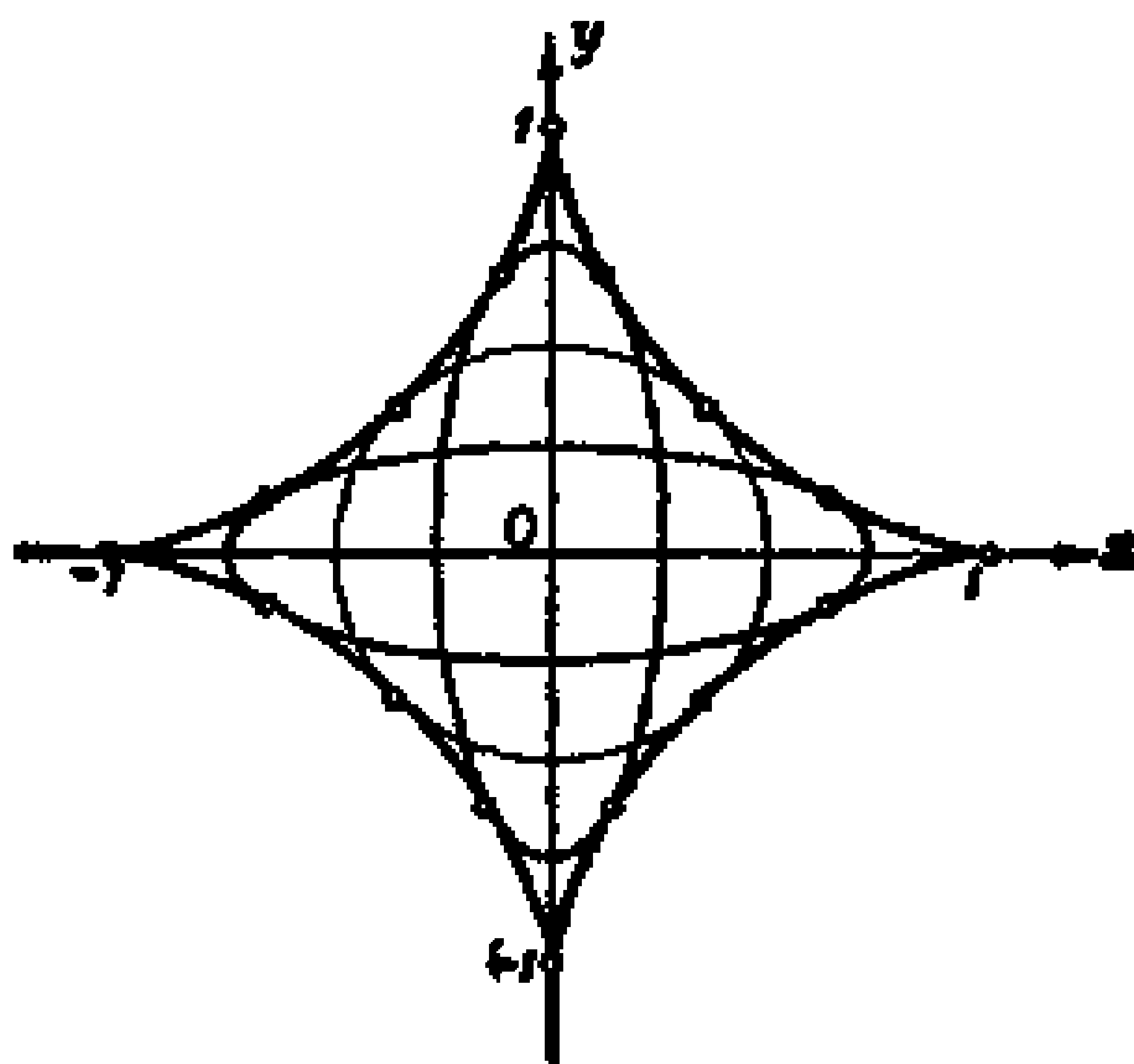


图 3.22 星形线作为椭圆族的包络

6. 星形线 $x^{2/3}+y^{2/3}=1$ 也作为椭圆族

$$\frac{x^2}{c^2}+\frac{y^2}{(1-c)^2}=1$$

的包络而出现, 它们的半轴 c 与 $(1-c)$ 的和为常量 1(图 3.22).

7. 曲线族 $(x-c)^2-y^3=0$ 表明, 在某些情况下我们找包络的办法可能失败. 在这里该法则给出的是 x 轴, 但是图 3.23 表明它不是包络, 它是曲线族的尖点的轨迹.

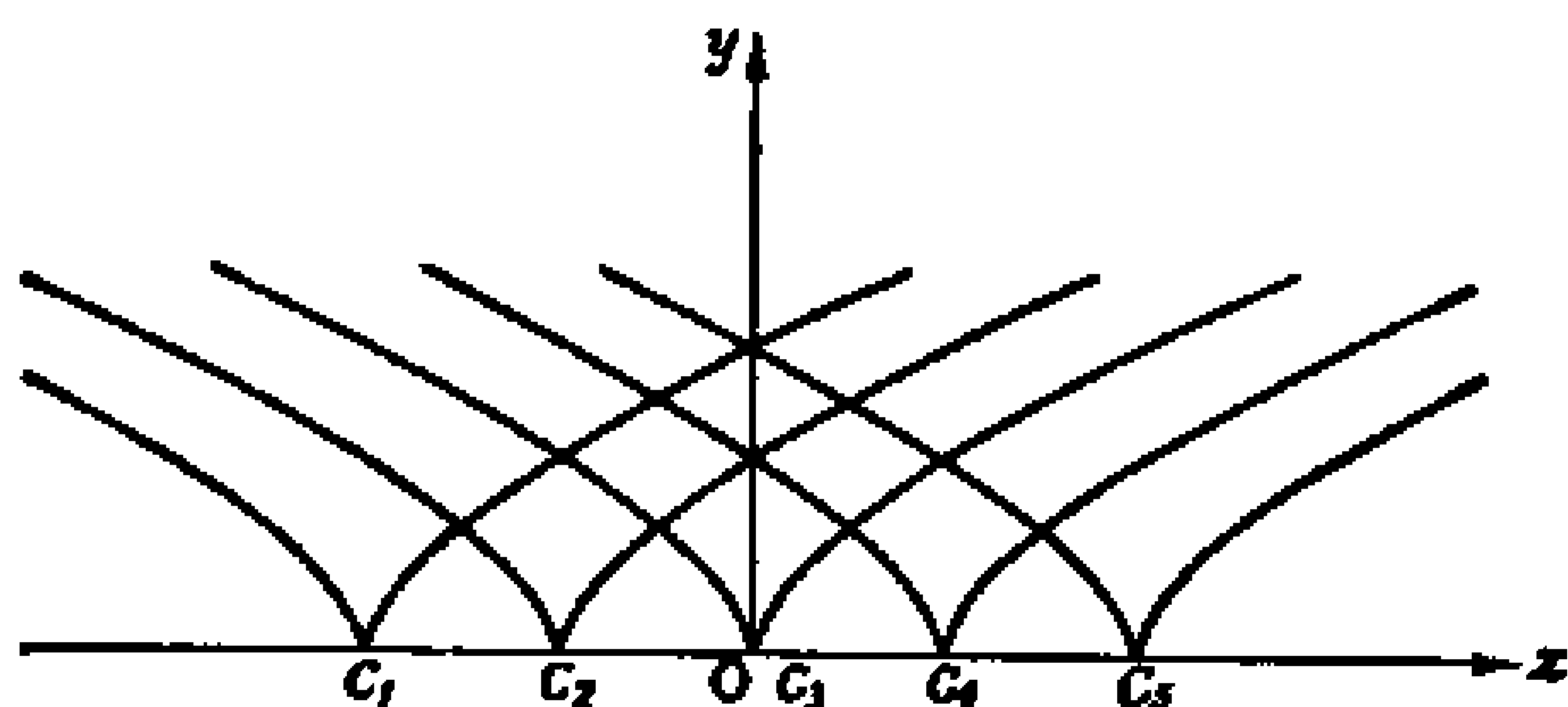


图 3.23　曲线族 $(x-c)^2-y^3=0$

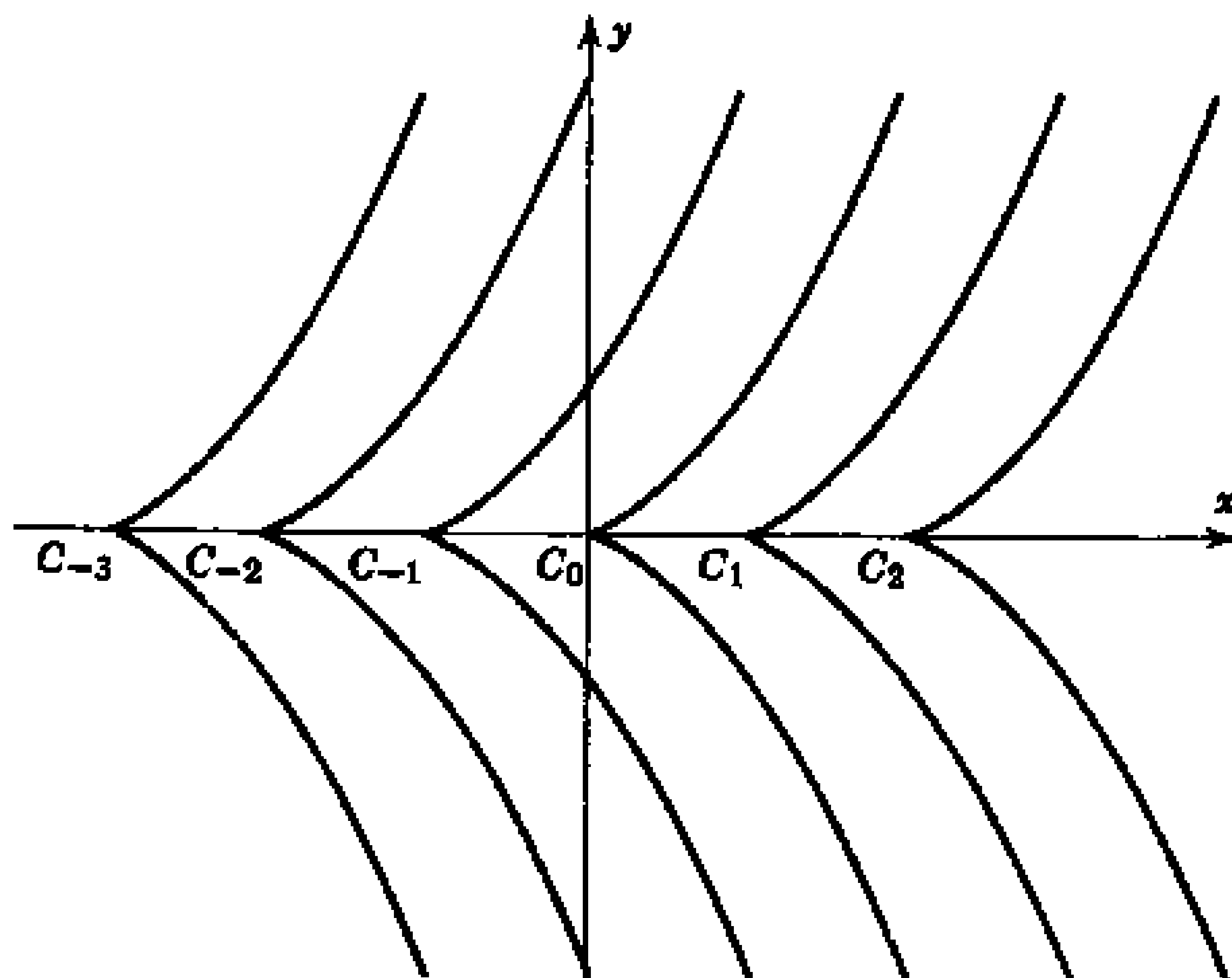


图 3.24　曲线族 $(x-c)^3-y^2=0$

8. 曲线族

$$(x-c)^3-y^2=0$$

的判别曲线是 x 轴 (见图 3.24). 这又是尖点的轨迹, 但是它与每条曲线相切, 因而在这种意义下必须看作是包络.

9. 环索曲线族

$$[x^2+(y-c)^2](x-2)+x=0$$

(见图 3.25) 具有以包络加重点轨迹所构成的判别曲线. 族中曲线彼此是全等的, 沿 y 轴方向平行移动可由一条曲线产生出另一条曲

线. 由微分法我们得到

$$f_c = -2(y-c)(x-2) = 0,$$

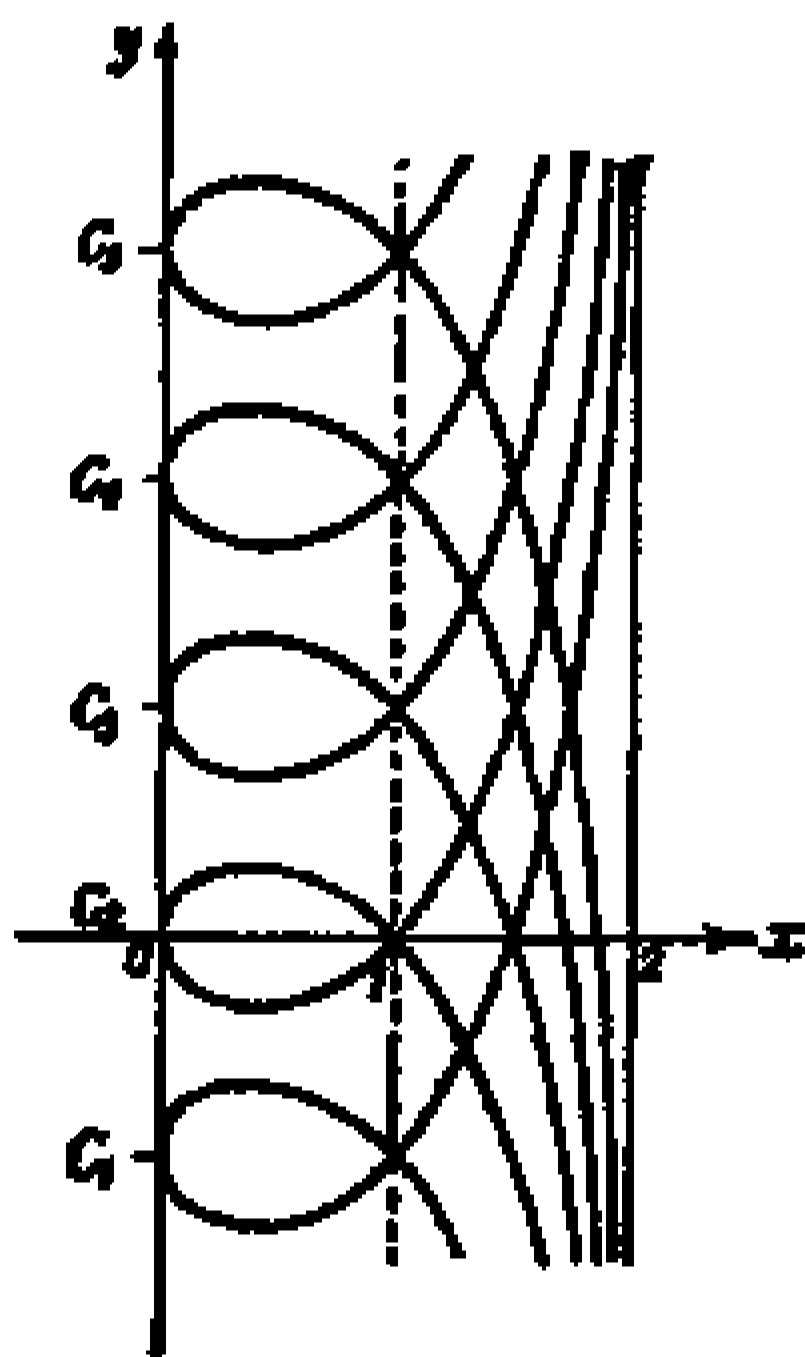


图 3.25　环索线族

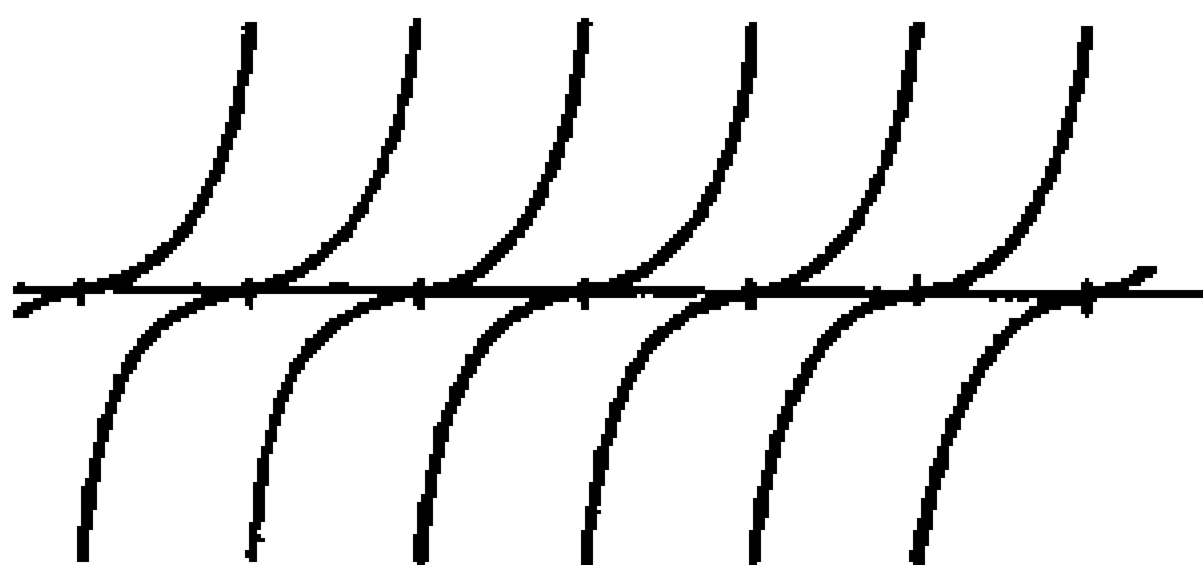


图 3.26　三次抛物线族

因而我们必有 $x=2$ 或者 $y=c$. 然而直线 $x=2$ 可置而不论, 因为没有 y 的有限值对应于 $x=2$. 因此我们有 $y=c$. 所以判别曲线是

$$x^2(x-2)+x=0.$$

这个曲线由直线 $x=0$ 和 $x=1$ 构成. 正像我们在图 3.25 中看到的, 只有 $x=0$ 是包络; 直线 $x=1$ 通过曲线族的全部重点.

10. 包络不必是相邻曲线交点的轨迹; 这由恒平行的三次抛物线族 $y-(x-c)^3=0$ 可以看出. 族中任何两条曲线皆不相交. 依上述法则给出方程 $f_c = 3(x-c)^2 = 0$, 因此 x 轴 $y=0$ 是判别曲线. 因族中所有曲线都与之相切, 所以 x 轴也是包络 (图 3.26).

11. 包络的概念使我们能对曲线 C 的**渐屈线**给出一个新的**定义**(见第一卷第 403 及 473 页以下). 设 C 由

$$x=\varphi(t), \qquad y=\psi(t)$$

给出. 我们定义 C 的渐屈线 E 作为 C 的法线的包络. 因为 C 的法线由

$$\{x-\varphi(t)\}\varphi'(t)+\{y-\psi(t)\}\psi'(t)=0$$

给出, 包络可由关于 t 微分这个方程而求得

$$0=\{x-\varphi(t)\}\varphi''(t)+\{y-\psi(t)\}\psi''(t)-\varphi'^2(t)-\psi'^2(t).$$

由这一方程和前一方程, 我们得到包络的参量表示

$$x=\varphi(t)-\psi'(t)\frac{\varphi'^2+\psi'^2}{\psi''\varphi'-\varphi''\psi'}=\varphi-\frac{\psi'\rho}{\sqrt{\varphi'^2+\psi'^2}},$$

$$y=\psi(t)+\varphi'(t)\frac{\varphi'^2+\psi'^2}{\psi''\varphi'-\varphi''\psi'}=\psi+\frac{\varphi'\rho}{\sqrt{\varphi'^2+\psi'^2}},$$

其中

$$\rho=\frac{(\varphi'^2+\psi'^2)^{3/2}}{\psi''\varphi'-\varphi''\psi'}$$

表示曲率半径 (见第一卷第 399 页). 这些方程完全与第一卷 (第 400 页) 给出的渐屈线方程相同.

12. 设曲线 C 由 $x=\varphi(t)$, $y=\psi(t)$ 给出. 我们来构造中心在 C 上且通过原点 O 的那些圆的包络 E. 因为诸圆由

$$x^2+y^2-2x\varphi(t)-2y\psi(t)=0$$

给出, 所以 E 的方程是

$$x\varphi'(t)+y\psi'(t)=0.$$

因此, 如果 P 是点 $(\varphi(t),\psi(t))$ 而 $Q(x,y)$ 是 E 的相应点, 则 OQ 垂直于 C 在 P 点的切线. 因为由定义 $PQ=PO$, 所以 PO 和 PQ 与 C 在点 P 的切线构成等角.

如果我们把 O 想像为一个发光点而 C 想像为一条反射曲线, 则 QP 是相应于 OP 的反射线. 反射线的包络叫做 C 关于 O 的焦

散曲线. 焦散曲线是 E 的渐屈线: 因为中心在 P 的圆与 E 在 Q 相切, 所以反射线 PQ 是 E 的法线, 而 E 的法线的包络就是它的渐屈线, 这正如我们在上例中已经看到的.

例如, 设 C 是一个通过 O 的圆, 则 E 是一个动圆 C' 上的点 O 的轨迹, 动圆 C' 与圆 C 全等, 且当 O' 与 O 重合时开始在圆 C 上滚动, 因为在滚动过程中, O 与 O' 始终关于这两圆的公共切线对称. 于是, E 将是一个特殊的外摆线, 事实上是一条心脏线. (见第一卷第 373 页) 因为外摆线的渐屈线是一条相似的外摆线 (见第一卷第 478 页), 所以在这种情况下, C 关于 O 的焦散曲线是一条心脏线.

练 习 3.5 c

1. 一个抛射体以初始倾角 α 和固定初速度 v 从原点射出, 沿着由方程

$$x = (v\cos\alpha)t,$$

$$y = (v\sin\alpha)t - \frac{1}{2}gt^2$$

所给出的抛物线轨道运行, 其中 g 是代表重力加速度的常数.

(a) 求以 α 为参量的轨道族的包络;

(b) 证明在该包络的上方没有点被击中;

(c) 证明在该包络的下方的每一点能以两种方式被击中, 也就是要证明每一个这样的点都位于两条抛射线上.

2. 求下列曲线族的包络:

(a) $y = cx + \dfrac{1}{c}$;

(b) $y^2 = c(x - c)$;

(c) $cx^2 + \dfrac{y^2}{c} = 1$;

(d) $(x - c)^2 + y^2 = a^2c^2/(1 + a^2)$, $a=$ 常数.

3. 设 C 是平面上的一条任意曲线，并考虑半径为 p 中心在 C 上的圆. 证明，这些圆的包络由两条平行于 C 而与 C 的距离为 p 的曲线组成.

4. 空间中的直线族可由依赖于参量 t 的两平面的交线给出

$$a(t)x + b(t)y + c(t)z = 1,$$

$$d(t)x + e(t)y + f(t)z = 1.$$

证明，如果这些直线是某曲线的切线 (即具有包络)，则

$$\begin{vmatrix} a-d & b-e & c-f \\ a' & b' & c' \\ d' & e' & f' \end{vmatrix} = 0.$$

5. 若平面曲线 C 是由 $x = f(t), y = g(t)$ 所给出，则它的配极曲线 C' 定义为直线族

$$\xi f(t) + \eta g(t) = 1$$

的包络，其中 (ξ, η) 是流动坐标.

(a) 证明， C 也是 C' 的配极曲线;

(b) 寻求圆 $(x-a)^2 + (y-b)^2 = 1$ 的配极曲线;

(c) 寻求椭圆 $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ 的配极曲线.

6. 一个半径为 a 的圆带着一条与它相对固定的切线在一条固定的直线上滚动. 取动切线与固定直线重合时的位置为极轴，证明，切线的包络由

$$x = a(\theta + \cos\theta\sin\theta - \sin\theta),$$

$$y = a(\cos^2\theta - \sin\theta)$$

给出，其中 θ 是动切线与极轴的夹角.

7. 在平面上，求通过固定点 O 的诸圆的包络，这些圆的圆心描过一条以 O 为中心的二次曲线.

8 (a) 如果 Γ 是一条平面曲线，而 O 是平面上一点，则 O 在 Γ 的变动切线上的正交投影的轨迹 Γ' 叫做 Γ 关于 O 点的垂足曲线. 证明，如果点 M 描过曲线 Γ, 则垂足曲线 Γ' 是以向径 OM 为直径的变圆的包络.

(b) 如果 Γ 是一个圆，而 O 是圆周上的一个点，则包络是什么样的？

9. 设 MM' 是平行于椭圆短轴的椭圆的变弦，求以 MM' 为直径的变圆的包络.

d. 曲面族的包络

上面关于曲线族包络的论述无需多大更动就可适用于曲面族. 给定一个确定在参量 c 的某区间上的单参量曲面族 $f(x,y,z,c)=0$, 我们称一张曲面 E 是该曲面族的包络，如果它和族中每一曲面沿一条整个曲线相切，并且，进一步，如果那些在 E 上的接触曲线形成一个完全覆盖 E 的单参量曲线族.

中心在 z 轴上，以 1 为半径的全部球面所形成的曲面族就是一例. 从直观上我们看出，包络是以 1 为半径，以 z 轴为其轴的圆柱面 $x^2+y^2-1=0$; 很简单，接触曲线族就是以 1 为半径，中心在 z 轴上，且平行于 xy 平面的圆族.

正像第 322 页上一样，如果我们假定包络确实存在，我们就能用下述的启发性的方法找出来：我们首先考虑相应于两个不同的参量 c 和 $c+h$ 的曲面 $f(x,y,z,c)=0$ 和 $f(x,y,z,c+h)=0$. 这两个方程确定了两个曲面的交线 (我们明确地假定这样一条交线存在). 作为上面两个方程的推论，这条曲线也满足这第三个方程

$$\frac{f(x,y,z,c+h)-f(x,y,z,c)}{h}=0.$$

如果我们让 h 趋向于 0, 交线将趋近于一个极限位置，并且极限曲线由这两个方程

$$f(x,y,z,c)=0, \qquad f_c(x,y,z,c)=0 \tag{54}$$

所确定. 在非严密的直观意义上, 这条曲线常被看作族中相邻曲面的交线. 它是参量 c 的函数, 所以对所有不同的 c 值, 交线在空间中形成一个单参量曲线族. 如果从上面的两个方程中消去量 c, 我们就得到一个方程, 称之为 **判别方程**. 正像在第 300 页一样, 我们能够证明, 包络一定满足这个判别方程.

正像在平面曲线的情况一样, 我们可以容易使我们自己相信, 只要 $f_x^2+f_y^2+f_z^2\neq 0$, 和判别曲面相切的平面也同时和族中相应的曲面相切. 因此, 判别曲面还给出了曲面族的包络和族中曲面的奇点的轨迹.

作为第一个例, 我们考虑上面提到的球面族

$$x^2+y^2+(z-c)^2-1=0.$$

为了求包络, 我们有附加方程

$$-2(z-c)=0.$$

对固定的 c 值, 这两个方程显然表示在高 $z=c$ 处平行于 xy 平面的以 1 为半径的圆. 如果在这两个方程之间消去参量 c, 我们就得到形如 $x^2+y^2-1=0$ 的包络的方程, 这是以 1 为半径, 以 z 轴为轴的直圆柱面的方程.

对于曲面族来说, 求双参量曲面族 $f(x,y,z,c_1,c_2)=0$ 的包络也是可能的. (但是, 对于曲线族而言, 包络的概念仅对单参量曲线族有意义.) 例如, 我们考虑所有以 1 为半径, 中心在 xy 平面上的球面, 其方程为

$$(x-c_1)^2+(y-c_2)^2+z^2-1=0.$$

直观立刻告诉我们, 两个平面 $z=1$ 和 $z=-1$ 在每一点和族中某一曲面相切. 在一般情况下, 我们说 E 是双参量曲面族的包络, 如果在 E 的每一点 P 曲面 E 与族中某一曲面以这样一种方式相切, 即当 P 在 E 上变化时, 相应于和 E 在 P 点相切的曲面的参量

c_1, c_2 的值在 c_1c_2 平面上某一区域内变化，而且，不同的点 (c_1, c_2) 相应于 E 上不同的点 P. 这时族中一曲面和包络在**一点**相切，而不是像前面那样沿一条整个曲线相切.

在类似于平面曲线情况中所做的假定之下，我们来求族中曲面与包络的接触点，如其存在，必满足方程

$$f(x,y,z,c_1,c_2)=0, \qquad f_{c_1}(x,y,z,c_1,c_2)=0,$$

$$f_{c_2}(x,y,z,c_1,c_2)=0.$$

指定参量的相应值，从这三个方程我们就确定出族中给定曲面的接触点. 反过来，如果我们消去参量 c_1, c_2，我们就得到包络必须满足的方程.

例如，以 1 为半径，中心在 xy 平面上的球面族，由具有两个参数 c_1 和 c_2 的方程

$$f(x,y,z,c_1,c_2)=(x-c_1)^2+(y-c_2)^2+z^2-1=0$$

给出. 构成包络的法则给出两个方程

$$f_{c_1}=-2(x-c_1)=0 \qquad \text{和} \qquad f_{c_2}=-2(y-c_2)=0.$$

于是，关于判别方程我们有 $z^2-1=0$，而事实上，正像我们从直观上已经看到的，平面 $z=1$ 和 $z=-1$ 就是包络.

练 习 3.5 d

1. 在 O 点有公共中心，轴平行于坐标轴的有定常体积 (即半轴的乘积是常数) 的椭球面族的包络是什么？

2. 平面族 $ax+by+cz=1$ 的包络是什么，其中 $a^2+b^2+c^2=1$?

3. (a) 求满足

$$OP+OQ+OR=\text{常数}=1$$

的双参量平面族的包络，这里 P,Q,R 表示平面与坐标的交点，而 O 是原点.

(b) 求满足

$$OP^2+OQ^2+OR^2=1$$

的平面族的包络.

4. 设平面族由

$$x\cos t+y\sin t+z=t$$

给定，其中 t 是参量.

(a) 在柱坐标 (r,θ,z) 中，求平面族包络的方程.

(b) 证明包络由某曲线的切线构成.

5. 设 $z=u(x,y)$ 是管面的方程，管面即中心在 xy 平面上的某条曲线 $y=f(x)$ 上而半径为 1 的球面族的包络. 证明 $u^2(u_x^2+u_y^2+1)=1$.

6. 求切于三个球面

$$S_1:\left(x-\frac{3}{2}\right)^2+y^2+z^2=\frac{9}{4},$$

$$S_2:x^2+\left(y-\frac{3}{2}\right)^2+z^2=\frac{9}{4},$$

$$S_3:x^2+y^2+\left(z-\frac{3}{2}\right)^2=\frac{9}{4}$$

的球面族的包络.

7. 设 Γ 是一条平面曲线， Γ' 是一条在第 311 页习题 8所述的垂足曲线.

(a) 设 M 描过曲线 Γ. 以向径 OM 作为直径的变球面的包络是什么？

(b) 如果 Γ 是一个圆， O 是圆周上的一个点，变球面的包络是什么？

8. 证明，曲面 $xyz=$ 常数是那种平面族的包络，它们和坐标面构成的四面体的体积是常数 (即截距的乘积是常数).

9. 设一平面在运动中总是保持切于抛物线 $z=0,\ y^2=4x$ 和 $y=0,\ z^2=4x$. 证明，它的包络由两个抛物柱面所构成.

3.6 交错微分型

a. 交错微分型的定义

在第一章 (第 86 页) 里，我们考虑过三个自变量的一般线性微分型

$$L=A(x,y,z)dx+B(x,y,z)dy+C(x,y,z)dz. \tag{55a}$$

沿任何有参量表示式 $x=\phi(t),\ y=\psi(t),\ z=\chi(t)$ 的曲线 Γ，微分型 L 决定了下式的值

$$\frac{L}{dt}=A\frac{dx}{dt}+B\frac{dy}{dt}+C\frac{dz}{dt}=A\dot{\phi}+B\dot{\psi}+C\dot{\chi}, \tag{55b}$$

它依赖于 Γ 的特定的参量表示. 如果 Γ 采用一个不同的参量 τ，我们就得到

$$\begin{aligned}\frac{L}{d\tau}&=A\frac{dx}{d\tau}+B\frac{dy}{d\tau}+C\frac{dz}{d\tau}\\&=\left(A\frac{dx}{dt}+B\frac{dy}{dt}+C\frac{dz}{dt}\right)\frac{dt}{d\tau}\\&=\frac{L}{dt}\frac{dt}{d\tau}.\end{aligned} \tag{55c}$$

然而，积分

$$\int_\Gamma L=\int\frac{L}{dt}dt=\int\left(A\frac{dx}{dt}+B\frac{dy}{dt}+C\frac{dz}{dt}\right)dt$$

只依赖于曲线 Γ(以及它的方向)，而不依赖于参量表示.

类似地，我们可以考虑微分型 ω，它是 dx,dy 与 dz 的二次型，即 ω 是符号 $dxdx,\ dxdy, dxdz, dydx, dydy,\ dydz, dzdx, dzdy, dzdz$ 的一个线性组合，其系数都是 x,y,z 的函数. 在空间的任何一个具有

参量表示 $x=\phi(s,t)$, $y=\psi(s,t)$, $z=\chi(s,t)$ 的曲面 S 上，微分型 ω 都定义了 $\omega/dsdt$ 的值，只要我们约定把商式

$$\frac{dxdx}{dsdt},\frac{dxdy}{dsdt},\frac{dxdz}{dsdt},\cdots$$

分别理解为雅可比行列式 [1]

$$\frac{d(x,x)}{d(s,t)},\frac{d(x,y)}{d(s,t)},\frac{d(x,z)}{d(s,t)},\cdots.$$

对于在曲面的每一点上具有同样的值 $\omega/dsdt$ 的两个微分型 ω, 我们是不加区别的. 由于行列式的交错特性，即

$$\frac{d(x,x)}{d(s,t)}=0,\quad \frac{d(x,y)}{d(s,t)}=-\frac{d(y,x)}{d(s,t)},\cdots,$$

我们看到，在 ω 的各项中含有 $dxdx, dydy, dzdz$ 的这些项没有贡献，而 $dydx, dzdy, dxdz$ 可以各自换成 $-dxdy, -dydz, -dzdx$. 因此， dx, dy, dz 的最普遍的二次微分型可以写成

$$\omega=a(x,y,z)dydz+b(x,y,z)dzdx+c(x,y,z)dxdy. \tag{56a}$$

微分型 ω 联系到一个曲面 S 的用参量 s,t 表示的点所给出的值是

$$\frac{\omega}{dsdt}=a(x,y,z)\frac{d(y,z)}{d(s,t)}+b(x,y,z)\times\frac{d(z,x)}{d(s,t)}+c(x,y,z)\frac{d(x,y)}{d(s,t)}. \tag{56b}$$

给 S 以不同的参量 s', t', 由雅可比行列式的乘法定律 (见第 282 页), 我们得到

$$\frac{\omega}{ds'dt'}=a\frac{d(y,z)}{d(s',t')}+b\frac{d(z,x)}{d(s',t')}+c\frac{d(x,y)}{d(s',t')}$$

---

1) 这个约定构成了**交错**微分型的特征. 在其他章节里，非交错的二次微分型也同样见到，如像空间或曲面上的线素的平方

$$ds^2=dx^2+dy^2+dz^2=Edu^2+2Fdudv+Gdv^2$$

所给出的微分型就是其中之一 (见第 310 页).

$$= \frac{\omega}{dsdt}\frac{d(s,t)}{d(s',t')}. \tag{56c}$$

稍后，我们还将定义二重积分

$$\iint_S \omega$$

并且将看到它不依赖于曲面 S 的特定的参量表示.

类似地，我们可以考虑 dx, dy, dz 的三次微分型 ω. 这样一个微分型对应于任何参量表示

$$x = \phi(r,s,t), y = \psi(r,s,t), z = \chi(r,s,t)$$

确定了 $\omega/drdsdt$ 的值，这里我们又把商式

$$\frac{dxdxdx}{drdsdt}, \frac{dxdydz}{drdsdt}, \cdots$$

理解为雅可比行列式

$$\frac{d(x,x,x)}{d(r,s,t)}, \frac{d(x,y,z)}{d(r,s,t)}, \cdots.$$

由于雅可比行列式的两个变量恒等时等于零，把两个变量互换时变号，因此三个自变量 x, y, z 的三次微分型全都是

$$\omega = a(x,y,z)dxdydz \tag{56d}$$

这样的类型. 当 x, y, z 是 r, s, t 的函数时，我们得到微分型的值

$$\frac{\omega}{drdsdt} = a(x,y,z)\frac{d(x,y,z)}{d(r,s,t)}. \tag{56e}$$

继续运用同样的方式进行下去，我们能够定义 dx, dy, dz 的 4 次，5 次，$\cdots$“交错”微分型. 但是所有这些微分型都恒等于零，这是因为我们所能组成的任何 4 阶，5 阶，$\cdots$ 雅可比行列式总有

两个因变量是恒等的，因而总是等于零[1]．

练 习 3.6 a

1. 试求下列各式的 $\omega/dudv$:

(a) $\omega = xdydz + ydzdx + zdxdy$,

$x = \cos u \sin v,\ y = \sin u \sin v,\ z = \cos v;$

(b) $\omega = (y - z)dydz + (z - x)dzdx + (x - y)dxdy$,

$x = au + bv,\ y = bu + cv,\ z = cu + av;$

(c) $\omega = dydz + dzdx + dxdy$,

$x = u^2 + v^2, y = 2uv, z = u^2 - v^2.$

b. 微分型的和与积

两个同次的微分型 (即两个线性的，两个二次的或两个三次的) 只要按对应的系数相加，就可以加起来，例如，对于

$$\omega_1 = a_1 dydz + b_1 dxdz + c_1 dxdy,$$

$$\omega_2 = a_2 dydz + b_2 dxdz + c_2 dxdy,$$

---

1) 高次型在高维空间中却有重要意义，在 4 维 (x, y, z, u) 空间中，最一般的 1,2,3,4 次交错微分型可以写成

$$Adx + Bdy + Cdz + Ddu, \tag{56f}$$

$$Adxdy + Bdydz + Cdzdu + Ddudx + Edxdz + Fdydu, \tag{56g}$$

$$Adydzdu + Bdzdudx + Cdudxdy + Ddxdydz, \tag{56h}$$

$$Adxdydzdu, \tag{56i}$$

其中系数 $A, B, \cdots$ 都是 x, y, z, u 的函数．这里，高于四次的微分型都等于零．

我们定义

$$\omega_1 + \omega_2 = (a_1 + a_2)dydz + (b_1 + b_2)dxdz + (c_1 + c_2)dxdy. \quad (57a)$$

我们可以定义任何两个同次的或不同次的微分型 ω_1 与 ω_2 的积 $\omega_1\omega_2$, 只要对于 ω_1 与 ω_2 代入它们的 dx, dy, dz 的表示式, 并引用乘法分配律, 但应注意, 在每一项中要保持微分原来的顺序[1]. 例如, 两个线性型

$$\omega_1 = A_1dx + B_1dy + C_1dz \text{ 与 } \omega_2 = A_2dx + B_2dy + C_2dz$$

的积是二次型

$$\begin{aligned}\omega_1\omega_2 &= (A_1dx + B_1dy + C_1dz)(A_2dx + B_2dy + C_2dz)\\ &= A_1A_2dxdx + A_1B_2dxdy + A_1C_2dxdz + B_1A_2dydx\\ &\quad + B_1B_2dydy + B_1C_2dydz + C_1A_2dzdx \qquad (57b)\\ &\quad + C_1B_2dzdy + C_1C_2dzdz\\ &= (B_1C_2 - C_1B_2)dydz + (C_1A_2 - A_1C_2)dzdx\\ &\quad + (A_1B_2 - B_1A_2)dxdy.\end{aligned}$$

如果我们用"系数向量" $R_1 = (A_1, B_1, C_1)$ 与 $R_2 = (A_2, B_2, C_2)$ 来描述微分型 ω_1 与 ω_2, 那么积 $\omega_1\omega_2$ 的系数恰好就是**向量积** $R_1 \times R_2$ 的分量 (见第 196 页). 很清楚, 微分型的积是不可交换的. 例如这里有 $\omega_1\omega_2 = -\omega_2\omega_1$.

一次型

$$\omega_1 = Adx + Bdy + Cdz$$

乘以二次型

$$\omega_2 = adydz + bdzdx + cdxdy,$$

1) 用这个方式组成的积有时候用符号 $\omega_1 \wedge \omega_2$ 表示.

我们就同样地得到

$$\omega_1\omega_2 = (Adx + Bdy + Cdz)(adydz + bdzdx + cdxdy)$$

$$= Aadxdydz + Abdxdzdx + Acdzdxdy$$

$$+ Badydydz + Bbdydzdx + Bcdydxdy$$

$$+ Cadzdydz + Cbdzdzdx + Ccdzdxdy$$

$$= (Aa + Bb + Cc)dxdydz. \qquad (57c)$$

我们看到，在这种情形 $\omega_1\omega_2$ 的系数是系数向量 (A,B,C) 与 (a,b,c) 的数量积. 这里碰巧有 $\omega_1\omega_2 = \omega_2\omega_1$.

作一个一次型与一个三次型，或两个二次型，或一个二次型与一个三次型的积，都得到高于三次的微分型，因而都等于零. 为了完整起见，把零次微分型定义为数量 $\alpha(x,y,z)$ 是方便的. 零次微分型 α 与任意次 $(k = 0,1,2,3)$ 的微分型 ω 的积就是用 d 乘 ω 的每一个系数.

容易由定义看出，微分型的乘法是适合结合律的. 例如对于以下三个线性型

$$L_i = A_idx + B_idy + C_idz \quad (i = 1,2,3)$$

将要求在习题 5 中去证明

$$L_1(L_2L_3) = \begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} dxdydz, \qquad (57d)$$

而对于 $(L_1L_2)L_3$，我们得到相同的算值.

当然，在自变量的个数多于三个的时候，就可以形成更多样的微分型的积.

练 习 3.6 b

1. 试计算下列的积：

(a) $(xdx+ydy)(xdx-ydy)$;

(b) $[(x^2+y^2)dx+2xydy][2xydx+(x^2-y^2)dy]$;

(c) $(adx+bdy)(adydz+bdzdx+cdxdy)$;

(d) $(dx+dy+dz)(dydz-dxdy)$.

2. 试证：对于 x,y,z 的任何一个一次型 ω 有 $\omega^2=0$.

3. 试证：对于三个变量的一次型 ω_1,ω_2, 有

$$(\omega_1+\omega_2)(\omega_1-\omega_2)=2\omega_2\omega_1.$$

4. 试证：对于三个变量的一次型，有

$$(\omega_1+\omega_2+\omega_3+\omega_4)(\omega_1-\omega_2+\omega_3-\omega_4)$$

$$=2(\omega_2+\omega_4)(\omega_1+\omega_3).$$

5. 试证 (57d).

c. 微分型的外微商

对于零次微分型，也就是对于数量 $\alpha(x,y,z)$, 根据定义，我们有

$$d\alpha=\alpha_x dx+\alpha_y dy+\alpha_z dz. \tag{58a}$$

这个微分型的系数恰好是第 222 页中所说的向量 $\mathrm{grad}\alpha$ 的分量. 更一般地，我们对于任何微分型 ω 都定义相应的**外微商** $d\omega$. 为此，我们把 ω 作为一个和式写出，其中的每一项是以数量因子为前导的微分 dx、dy,dz 的一个乘积，并且把每个数量因子换成它的通常意义下的微分，这样来组成外微商. 例如，对于一次型

$$L=Adx+Bdy+Cdz,$$

我们求出对于 dL 的二次微分型

$$dL=dAdx+dBdy+dCdz=(A_xdx+A_ydy+A_zdz)dx$$

$$+ (B_x dx + B_y dy + B_z dz)dy + (C_x dx + C_y dy + C_z dz)dz \tag{58b}$$

$$= (C_y - B_z)dydz + (A_z - C_x)dzdx + (B_x - A_y)dxdy.$$

如果我们把向量 $R = (A, B, C)$ 与 L 联系起来，我们就有这样一个值得注意的事实： dL **的系数恰好是** R **的旋度的分量**(见第 226 页).

二次型

$$\omega = adydz + bdzdx + cdxdy$$

的外微商 $d\omega$ 是三次型

$$d\omega = dadydz + dbdzdx + dcdxdy$$

$$= (a_x dx + a_y dy + a_z dz)dydz + (b_x dx + b_y dy + b_z dz)dzdx \tag{58c}$$

$$+ (c_x dx + c_y dy + c_z dz)dxdy = (a_x + b_y + c_z)dxdydz.$$

因此，如果 ω 的系数组成向量 $R = (a, b, c)$，那么 $d\omega$ 的系数就是数量 $\text{div}R$ (见第 225 页).

三次微分型的微商是四次的，因而等于零.

一个重要的一般法则 (“庞加莱引理”) 是，**任何微分型的二次外微商等于零**

$$dd\omega = 0. \tag{58d}$$

在三维空间中，这只须对于 ω 是零次或一次的情形作出证明. 现在，如果 ω 是一个数量 $\alpha(x, y, z)$，根据 (58a,b) 我们有

$$d^2\omega = d(\alpha_x dx + \alpha_y dy + \alpha_z dz) = 0.$$

这实际上只是第 227 页上所叙述的法则的一种不同的表述方式，那里的说法是，对于任何一个数量 α 都有 $\text{curl}\,(\text{grad}\,\alpha) = 0$. 类似地，根据 (58b,c)，对于一次微分型

$$\omega = Adx + Bdy + Cdz$$

的情形，我们得到

$$d^2\omega = d[(C_y - B_z)dydz + (A_z - C_x)dzdx + (B_x - A_y)dxdy] = 0.$$

这也不是什么新的事实，只是对于任何向量 R，法则 $\operatorname{div}(\operatorname{curl} R) = 0$ 有效 (见第 227 页).

求以已知型 ω 为其外微商的型 τ，这个反问题是基本的，我们希望用一个适当的微分型 τ 把一个已知微分型 ω 表示成

$$\omega = d\tau. \tag{58e}$$

当这样一个表示法是可能的时候，我们称 ω 是一个**恰当微分**或**全微分**. 对微分 τ 应用法则 (59)，我们看到，ω **是一个恰当微分的一个必要条件是** $d\omega = 0$.[1] 其实这个条件也是充分的，就是说对于 $d\omega = 0$，倘若我们限制在 ω 的定义域[2]内一点 (x_0, y_0, z_0) 的一个长方体邻域内，方程 (58e) 就有一个解 τ.

我们分别对 ω 的每一个次数来证明这个陈述. 如果 ω 是一次的，设

$$\omega = Adx + Bdy + Cdz,$$

那么，根据 (58b)，条件 $d\omega = 0$ 等价于关系

$$C_y - B_z = 0, \quad A_z - C_x = 0, \quad B_x - A_y = 0. \tag{58f}$$

但是这恰好是**可积条件**，它使我们能够把 ω 表示成某个函数 f 的全微分，只要我们把点 (x, y, z) 限制在包含 (x_0, y_0, z_0) 点的一个长方体内，或者 (更一般地) 一个单连通的集合内 (见第 111 页).

对于二次的 ω,

$$\omega = adydz + bdzdx + cdxdy,$$

---

1) 满足 $d\omega = 0$ 的型 ω 称为闭的.

2) 我们总是假定这里所考虑的微分型具有使我们论证成立所需要的次数的连续的微商的系数.

则由 (58c) 可见，条件 $d\omega=0$ 等价于

$$a_x+b_y+c_z=0. \tag{58g}$$

假设这个条件在长方体

$$|x-x_0|<r_1, |y-y_0|<r_2, |z-z_0|<r_3$$

内满足，我们要证明的是 $\omega=d\tau$, 其中 τ 有这样的形式

$$\tau=Adx+Bdy+Cdz.$$

这意味着要求出函数 A,B,C 使得

$$a=C_y-B_z,\ b=A_z-C_x,\ c=B_x-A_y.$$

我们试图选取 $C\equiv0$ 来满足这些方程. 那么为了满足前两个方程, A 与 B 就必须具有这样的形式

$$A(x,y,z)=\alpha(x,y)+\int_{z_0}^{z}b(x,y,\zeta)d\zeta,$$

$$B(x,y,z)=\beta(x,y)-\int_{z_0}^{z}a(x,y,\zeta)d\zeta.$$

用条件 (58g), 就得到

$$\begin{aligned}\frac{\partial}{\partial z}(B_x-A_y)&=\frac{\partial}{\partial x}B_z-\frac{\partial}{\partial y}A\\&=-a_x-b_y=c_z.\end{aligned}$$

因此 B_x-A_y-c 不依赖于 z. 如果当 $z=z_0$ 时第三个方程 $c=B_x-A_y$ 成立，那么它就对所讨论的全部 z 都成立. 因此我们只要这样来确定函数 $\alpha(x,y)$ 与 $\beta(x,y)$, 使得

$$\beta_x(x,y)-\alpha_y(x,y)=c(x,y,z_0).$$

例如，取

$$\alpha(x,y)=0,\quad \beta(x,y)=\int_{x_0}^{x}c(\xi,y,z_0)d\xi$$

就成了.

最后，对于三次算子

$$\omega = \alpha(x,y,z)dxdydz,$$

条件 $d\omega = 0$ 总是满足的. 我们要把 ω 表示成 $\omega = d\tau$, 其中 τ 是一个二次微分型:

$$\tau = adydz + bdzdx + cdxdy.$$

根据 (58c), 这等于要求出函数 a, b, c, 使得

$$a_x + b_y + c_z = \alpha.$$

很明显，这里有一个解是

$$a(x,y,z) = b(x,y,z) = 0,$$

$$c(x,y,z) = \int_{z_0}^{z} \alpha(x,y,\zeta)d\zeta.$$

这就证明了我们的定理.

练　习　3.6 c

1. 试算出下列各式的 $d\omega$:

(a)$\omega = \operatorname{arctg} \frac{y}{x}$;

(b)$\omega = ydx - xdy$;

(c)$\omega = f(x,y)dxdy$;

(d)$\omega = x^2 \cos y \sin zdydz - x \sin y \sin zdzdx + x \cos zdxdy$;

(e)$\omega = (z^2 - y^2)xdydz + (x^2 - z^2)ydzdx + (y^2 - x^2)zdxdy$.

2. 试证：对于三个变量的一次型有

$$d(\omega_1\omega_2)=\omega_1 d\omega_2+(d\omega_1)\omega_2.$$

3. 试证：三个变量的恰当一次型的积是恰当的.

d. 任意坐标系中的外微分型

直到现在，我们总是把微分型看作是空间笛卡儿坐标 x, y, z 的微分 dx, dy, dz 的交错积的线性组合. 在定义两个微分型的积与一个微分型的微商时，我们所使用的主要就是 dx, dy, dz 的型的这种表示法，但是在应用中，交错微分型的用处依赖于这样一个事实：当三维欧氏空间使用的是任何一个曲线坐标 (u, v, w) 的时候，这些型是可以定义的，而且可以用相同的方式进行运算. 更一般地，这在任何一种非欧三维空间或使用参量 u, v, w 的流形[1]上 (例如四维欧氏空间中的三维"曲面"上) 也成立，重要的是，型的运算法能够通过一种与空间的坐标系无关的"不变方法"来定义，而且得到的公式在每个坐标系中看起来都相同.

在这部分论述中，人们把三维空间或流形 Σ 上的点 P 看作是独立于任何坐标系而存在的几何对象. 一个数量 f 是点 P 的一个实值函数 (也就是 Σ 到实数轴的一个映象). 有很多用曲线坐标描述 P 点的方法，也就是用三元数组 (u, v, w)，例如，用欧氏空间中的直角坐标或球坐标. 我们总是假定，任何两种这样的坐标系，如 u, v, w 与 u', v', w' 都是用变换方程式

$$u'=\phi(u, v, w),\ v'=\psi(u, v, w),\ w'=\chi(u, v, w)$$

联系起来的，其中 ϕ, ψ, χ 是连续函数，具有我们运算中所需要用到的各次连续微商，并且雅可比行列式[1]

$$\frac{d(u', v', w')}{d(u, v, w)}\neq 0.$$

---

1) 一般我们用"流形"这个术语表示 n 维欧氏空间中用参量形式给出的任何 m 维集合 $(m\leq n)$.

在这个条件下，u,v,w 可以用 u',v',w' 的类似的公式来表示. 在给定的 (u,v,w) 坐标系中，一个数量 $f=f(P)$ 就是 P 点坐标 u,v,w 的一个函数 $f(u,v,w)$. 在不同的坐标系中，表示同一数量的函数一般说来是很不同的.

设 C 是流形 Σ 上用参量形式 $P=P(t)$ 表示的一条曲线；对某个区间内的每一个实数 t, 参量方程联系到流形 Σ 上的一个点 P. 任何一个定义在 Σ 上的数量 $f(P)$, 通过复合式 $f(P(t))$ 在曲线 C 上产生 t 的一个函数. 如果这个函数是可微的，微商 $\frac{df}{dt}$ 就有意义，这个微商是对于曲线 C 及其参量表示而定义的，不依赖于流形 Σ 所用的曲线坐标. 在一个给定的坐标系中，P 点的坐标 u,v,w 本身都是函数 $u=u(t)$, $v=v(t)$, $w=w(t)$; 而 $f(P(t))$ 是由复合函数 $f(u(t),v(t),w(t))$ 给出的. 假定 $f(u,v,w)$ 与 $u(t),v(t),w(t)$ 有连续的微商，我们由微分法的链式规则求得 $\frac{df}{dt}$ 在特定的 u,v,w 坐标系中所取的形式为

$$\frac{df}{dt}=\frac{\partial f}{\partial u}\frac{du}{dt}+\frac{\partial f}{\partial v}\frac{dv}{dt}+\frac{\partial f}{\partial w}\frac{dw}{dt}. \tag{59}$$

在 Σ 内的一个零次微分型是一个数量 f. 一般的一次微分型 ω 定义为这样一个形式上的表达式

$$\omega=\sum_{i=1}^{N}a_i df_i,$$

其中 $a_1,a_2,\cdots,a_N$; $f_1,f_2,\cdots,f_N$ 是已知的数量. 沿任何一条用参量 t 表示的曲线 C, 我们把 ω 与 t 的函数联系起来，这个函数记作 $\frac{\omega}{dt}$, 它用下式来定义

$$\frac{\omega}{dt}=\sum_{i=1}^{N}a_i\frac{df_i}{dt}.$$

两个型

$$\omega=\sum_{i=1}^{N}a_i df_i \quad 与 \quad \omega'=\sum_{i=1}^{m}b_i dg_i$$

1) 变换所涉及到的单值函数 ϕ,ψ,χ 的特定表示只需在局部上成立，也就是在某个点的一个充分小的邻域内成立.

如果对于任何曲线 C 与沿 C 的参量 t 都有

$$\frac{\omega}{dt}=\frac{\omega'}{dt},$$

那么就认为是相等的.

在一个特定的 u,v,w 坐标系中 ω/dt 成为

$$\begin{aligned}\frac{\omega}{dt}&=\sum_{i=1}^{N}a_i\left(\frac{\partial f_i}{\partial u}\frac{du}{dt}+\frac{\partial f_i}{\partial v}\frac{dv}{dt}+\frac{\partial f_i}{\partial w}\frac{dw}{dt}\right)\\&=A\frac{du}{dt}+B\frac{dv}{dt}+C\frac{dw}{dt},\end{aligned}$$

其中

$$A=\sum_{i=1}^{N}a_i\frac{\partial f_i}{\partial u},\quad B=\sum_{i=1}^{N}a_i\frac{\partial f_i}{\partial v},\ C=\sum_{i=1}^{N}a_i\frac{\partial f_i}{\partial w}$$

是定义在 Σ 上的数量. 根据我们的一次微分型相等的定义, 可以把 ω 写成

$$\omega=Adu+Bdv+Cdw.$$

这里用特定的坐标系 u,v,w 表示的 ω 的系数 A,B,C 是唯一确定的; 因为如果我们把"坐标线"$u=t$, $v=$ 常数, $w=$ 常数, 取作曲线 C, 我们就得到

$$\frac{\omega}{dt}=\frac{\omega}{du}=A$$

并且同理有

$$\frac{\omega}{dv}=B,\quad \frac{\omega}{dw}=C.$$

因此, 在任何特定的坐标系 u,v,w 中, 我们可以把 ω 写成

$$\omega=\frac{\omega}{du}du+\frac{\omega}{dv}dv+\frac{\omega}{dw}dw, \tag{60}$$

其中 $\frac{\omega}{du}$ 实际上是代表沿 v,w 为常数的一条曲线上的偏微商. 这个公式可以认为是链式法则 (59) 从数量函数 f 的微分 df 到一般的一次微分型 ω 的一个推广.

我们可以用完全同样的方法把**二次交错微分型** ω 定义为这样一个形式上的表达式

$$\omega = \sum_{i=1}^{N} a_i df_i dg_i, \tag{61a}$$

其中 $a_1,\cdots,a_N$; $f_1,\cdots,f_N$; $g_1,\cdots,g_N$ 是定义在 Σ 上的数量. 在 Σ 内任何一个用参量 s,t 表示的曲面 S 上, 我们把 ω 联系到 $\frac{\omega}{dsdt}$ 的这个值

$$\frac{\omega}{dsdt} = \sum_{i=1}^{N} a_i \frac{d(f_i,g_i)}{d(s,t)} = \sum_{i=1}^{N} a_i \begin{vmatrix} \frac{\partial f_i}{\partial s} & \frac{\partial f_i}{\partial t} \\ \frac{\partial g_i}{\partial s} & \frac{\partial g_i}{\partial t} \end{vmatrix}. \tag{61b}$$

两个型 ω 与 ω', 尽管它们是用不同的数量函数来表示的, 但是只要它们在每一张曲面上对于每一种参数表示有相同的值

$$\frac{\omega}{dsdt} = \frac{\omega'}{dsdt},$$

就认为是恒等的. 现在在任何一个特定的坐标系 (u,v,w) 中, 对于两个数量 f 与 g, 我们有

$$\begin{aligned}\begin{vmatrix} f_s & f_t \\ g_s & g_t \end{vmatrix} &= \begin{vmatrix} f_u u_s + f_v v_s + f_w w_s & f_u u_t + f_v v_t + f_w w_t \\ g_u u_s + g_v v_s + g_w w_s & g_u u_t + g_v v_t + g_w w_t \end{vmatrix} \\ &= (f_v g_w - f_w g_v)(v_s w_t - v_t w_s) + (f_w g_u - f_u g_w)(w_s u_t - w_t u_s) \\ &\quad + (f_u g_v - f_v g_u)(u_s v_t - u_t v_s),\end{aligned}$$

所以

$$\frac{\omega}{dsdt} = a\frac{d(v,w)}{d(s,t)} + b\frac{d(w,u)}{d(s,t)} + c\frac{d(u,v)}{d(s,t)}, \tag{61c}$$

其中

$$a = \sum_{i=1}^{N} a_i \frac{d(f_i,g_i)}{d(v,w)}, \quad b = \sum_{i=1}^{N} a_i \frac{d(f_i,g_i)}{d(w,u)},$$

$$c = \sum_{i=1}^{N} a_i \frac{d(f_i, g_i)}{d(u, v)}. \tag{61d}$$

于是，我们在 (u, v, w) 系中可以把 ω 写成

$$\omega = a dv dw + b dw du + c du dv. \tag{61e}$$

在 ω 的这个表示式中的系数 a, b, c 仍然是唯一确定的；它们由下式给出

$$a = \frac{\omega}{dv dw}, \quad b = \frac{\omega}{dw du}, \quad c = \frac{\omega}{du dv},$$

其中 $a = \dfrac{\omega}{dv dw}$ 的值是对于坐标面 $v = s$, $w = t$, $u =$ 常数而计算的，b 与 c 类似地计算. 在 (u, v, w) 系中，ω 的符号表示式 (61c) 变成

$$\omega = \frac{\omega}{dv dw} dv dw + \frac{\omega}{dw du} dw du + \frac{\omega}{du dv} du dv, \tag{61f}$$

这与一次微分型的公式 (60) 相类似.[1]

在一张用参量 s, t 表示的曲面上，我们把两个一次型

$$L = \sum_i a_i df_i, \qquad M = \sum_k b_k dg_k \tag{62a}$$

的积 LM 定义为二次型 ω, 对于它有

$$\frac{\omega}{ds dt} = \frac{L}{ds}\frac{M}{dt} - \frac{L}{dt}\frac{M}{ds}$$

---

1) 公式 (61a, b) 对于使用参量 $u_1, u_2, \cdots, u_n$ 的 n 维空间中的二次型仍然有效，而代替 (61c,d,e,f), 我们有

$$\omega = \sum_{\substack{j,k=1,2,\cdots,n \\ j<k}} A_{jk} du_j du_k, \tag{61g}$$

其中

$$A_{jk} = \sum_i a_i \frac{d(f_i, g_i)}{d(u_j, u_k)} = \frac{\omega}{du_j du_k} \tag{61h}$$

是容易证实的.

$$= \sum_i a_i \frac{\partial f_i}{\partial s} \sum_k b_k \frac{\partial g_k}{\partial t} - \sum_i a_i \frac{\partial f_i}{\partial t} \sum_k b_k \frac{\partial g_k}{\partial s}$$

$$= \sum_{i,k} a_i b_k \frac{d(f_i, g_k)}{d(s,t)}^{1)}. \tag{62b}$$

由此，如果 L 与 M 由 (62a) 给出，LM 就可以等同于二次型

$$\omega = \sum_{i,k} a_i b_k df_i dg_k. \tag{62c}$$

然而由 (62b) 给出的

$$\frac{\omega}{dsdt} = \frac{LM}{dsdt}$$

的定义是与通过数量 a_i, f_i, b_k, g_k 表示的 L 与 M 的特定的表示法无关的，因此公式 (62c) 对于因子 L 与 M 的所有的表示法都一定表示同一个型 $\omega = LM$.

由一次型产生二次型的另一个方法是微分法，已知一次型

$$L = \sum_i a_i df_i, \tag{63a}$$

我们可以不用任何特定的坐标系，而用下述规定来定义 dL:

$$\begin{aligned}\frac{dL}{dsdt} &= \frac{\partial}{\partial s}\frac{L}{dt} - \frac{\partial}{\partial t}\frac{L}{ds} \\ &= \frac{\partial}{\partial s}\sum_i a_i \frac{\partial f_i}{\partial t} - \frac{\partial}{\partial t}\sum_i a_i \frac{\partial f_i}{\partial s} \\ &= \sum_i \left(\frac{\partial a_i}{\partial s}\frac{\partial f_i}{\partial t} - \frac{\partial a_i}{\partial t}\frac{\partial f_i}{\partial s}\right) \\ &= \sum_i \frac{d(a_i, f_i)}{d(s,t)}.\end{aligned} \tag{63b}$$

1) 这里像往常一样 M/ds, M/dt 分别表示对于 s, t 的"偏"微商("常"与"偏"微分法之间的一贯的区别可以不十分注意).

这等价于公式

$$dL = \sum_i da_i df_i, \tag{63c}$$

而且表明二次型 dL 与用数量 a_i, f_i 的 L 的特定表示法 (63a) 无关. 这是在型 L 表示成 $L = Adx + Bdy + Cdz$ 的特殊情况下, L 的微商公式 (58b) 的自然推广.

在一种特殊情况下, 即当一次型 L 是一个全微分 (也就是 $L = df$, 其中 f 是一个数量) 的时候, 我们从 (63c) 当然得到 $dL = 0$. 因此, 对于零次算子 f, 下述法则有效

$$ddf = 0.$$

当 L 在特定的空间坐标系 (u, v, w) 中表示成标准形式

$$L = Adu + Bdv + Cdw$$

时, 我们从 (61f) 和 (63b) 求得

$$\begin{aligned}
dL &= dAdu + dBdv + dCdw \\
&= \frac{dL}{dvdw}dvdw + \frac{dL}{dwdu}dwdu + \frac{dL}{dudv}dudv \\
&= \left(\frac{\partial}{\partial v}\frac{L}{dw} - \frac{\partial}{\partial w}\frac{L}{dv}\right)dvdw \\
&\quad + \left(\frac{\partial}{\partial w}\frac{L}{du} - \frac{\partial}{\partial u}\frac{L}{dw}\right)dwdv + \left(\frac{\partial}{\partial u}\frac{L}{dv} - \frac{\partial}{\partial v}\frac{L}{du}\right)dudv \\
&= (C_v - B_w)dvdw + (A_w - C_u)dwdu + (B_u - A_v)dudv,
\end{aligned}$$

与公式 (58b) 相符合.

如果 $dL = 0$, 由上所述我们得到

$$C_v - B_w = A_w - C_u = B_u - A_v = 0.$$

这说明, 在局部上存在一个数量 f, 使得 $A = f_u$, $B = f_v$, $C = f_w$ 或 $L = df$.

最后，一个三次交错微分型是用形式上的表达式

$$\omega = \sum_{i=1}^{N} a_i df_i dg_i dh_i \tag{64a}$$

定义，其中 a_i, f_i, g_i, h_i 是数量. 在任何一个空间坐标系 (r, s, t) 中，它决定下式的值

$$\frac{\omega}{drdsdt} = \sum_{i=1}^{N} a_i \frac{d(f_i, g_i, h_i)}{d(r, s, t)}. \tag{64b}$$

通过特定的 (u, v, w) 坐标系，我们可以写成

$$\frac{\omega}{drdsdt} = \sum_{i=1}^{N} a_i \frac{d(f_i, g_i, h_i)}{d(u, v, w)} \frac{d(u, v, w)}{d(r, s, t)} \tag{64c}$$

这等价于恒等式

$$\omega = adudvdw, \tag{64d}$$

其中

$$a = \sum_{i=1}^{N} a_i \frac{d(f_i, g_i, h_i)}{d(u, v, w)}. \tag{64e}$$ [1]

我们可以定义一次型

$$L = \sum_i a_i df_i$$

---

1) 在使用参量 $u_1, \cdots, u_n$ 的 n 维空间中，代替 (64c,d,e) 我们有公式

$$\omega = \sum_{\substack{j,k,m=1,\cdots,n \\ j<k<m}} A_{jkm} du_j du_k du_m,$$

其中

$$A_{j,k,m} = \sum_i a_i \frac{d(f_i, g_i, h_i)}{d(u_j, u_k, u_m)} = \frac{\omega}{du_j du_k du_m}.$$

与二次型

$$\omega = \sum_k b_k dg_k dh_k$$

的积 $L\omega$ 为

$$\frac{L\omega}{drdsdt} = \frac{L}{dr}\frac{\omega}{dsdt} + \frac{L}{ds}\frac{\omega}{dtdr} + \frac{L}{dt}\frac{\omega}{drds}$$

$$= \sum_{i,k} a_i b_k \left(\frac{\partial f_i}{\partial r}\frac{d(g_k, h_k)}{d(s,t)} + \frac{\partial f_i}{\partial s}\frac{d(g_k, h_k)}{d(t,r)} + \frac{\partial f_i}{\partial t}\frac{d(g_k, h_k)}{d(r,s)} \right)$$

$$= \sum_{i,k} a_i b_k \frac{d(f_i, g_k, h_k)}{d(r,s,t)}.$$

这等价于公式

$$L\omega = \sum_{i,k} a_i b_k df_i dg_k dh_k. \tag{65a}$$

正像可以从 L 与 ω 的表达式的形式上的乘法所能预期的一样，当 L 与 ω 在一个给定的 (u, v, w) 坐标系中取标准形式

$$L = Adu + Bdv + Cdw,$$

$$\omega = advdw + bdwdu + Cdudv$$

时，其积就变成

$$L\omega = (Aa + Bb + Cc)dudvdw, \tag{65b}$$

这与 (57c) 相符合.

二次微分型

$$\omega = \Sigma a_i dg_i dh_i$$

的微商，可以不依赖于特定的坐标系而用下述法则来定义:

$$\frac{d\omega}{drdsdt} = \frac{\partial}{\partial r}\frac{\omega}{dsdt} + \frac{\partial}{\partial s}\frac{\omega}{dtdr} + \frac{\partial}{\partial t}\frac{\omega}{drds}$$

$$= \frac{\partial}{\partial r}\sum_i a_i \frac{d(g_i, h_i)}{d(s, t)} + \frac{\partial}{\partial s}\sum_i a_i \frac{d(g_i, h_i)}{d(t, r)}$$

$$+ \frac{\partial}{\partial t}\sum_i a_i \frac{d(g_i, h_i)}{d(r, s)}.$$

这样，容易验证

$$\frac{d\omega}{drdsdt} = \sum_i \frac{d(a_i, g_i, h_i)}{d(r, s, t)}. \tag{66a}$$

所以，我们的 $d\omega$ 的定义隐含着

$$d\omega = \sum_i da_i dg_i dh_i. \tag{66b}$$

对取标准形式的 ω

$$\omega = adv dw + bdw du + cdu dv \tag{66c}$$

我们得到

$$d\omega = (a_u + b_v + c_w) du dv dw. \tag{66d}$$

这个关于 $d\omega$ 的特殊的表示式可以再一次用来(像第 344 页)证明：具有 $d\omega = 0$ 的二次微分型 ω 是可以局部地表示成 $\omega = dL$ 的，其中 L 是一个适当的一次微分型.

练 习 3.6 d

1. 在球坐标系 $x = \rho\sin\phi\cos\theta$, $y = \rho\sin\phi\sin\theta$, $z = \rho\cos\phi$ 中，取三个单位向量 $\mathbf{u}, \mathbf{v}, \mathbf{w}$，它们的方向分别是 r, ϕ, θ 诸线的方向. 试证：

$$d\mathbf{X} = (dx, dy, dz) = \mathbf{u}d\rho + \mathbf{v}\rho d\phi + \mathbf{w}\rho\sin\phi d\theta.$$

从面在球坐标系中求出 $\nabla f(\rho, \phi, \theta)$ 的表达式，其中 ∇f 通过

$$\nabla f \cdot d\mathbf{X} = df$$

来定义.

3.7 最大与最小

a. 必要条件

对于多元函数, 像对于单元函数一样, 微分法的一个非常重要的应用是最大与最小的理论.

我们从考虑两个自变量 x, y 的函数 $u = f(x, y)$ 开始. 函数的定义域是 xy 平面内的某个集合 R. 我们可以在 xyz 空间内用方程 $z = f(x, y)$ 所确定的曲面 S 来表示函数 f. 我们说 $f(x, y)$ 在它的定义域 R 内的 (x_0, y_0) 点有一个**最大**[1), 如果对于 R 内所有的点 (x, y) 都有 $f(x_0, y_0) \geq f(x, y)$. 这样的一个最大值对应于曲面 S 上的一个最高点. 我们说**严格**最大值, 如果对于 R 内一切不同于 (x_0, y_0) 的点 (x, y) 都有 $f(x_0, y_0) > f(x, y)$, 从而函数的最大值只在一个点 (x_0, y_0) 达到. 类似地, 如果对于 R 内所有的点 (x, y) 都有 $f(x_1, y_1) \leq f(x, y)$, 就说 $f(x, y)$ 在 (x_1, y_1) 点有一个**最小**, 如果对于 R 内所有的点 $(x, y) \neq (x_1, y_1)$ 都有 $f(x_1, y_1) < f(x, y)$, 就说 $f(x, y)$ 有一个**严格最小**. 第 121 页的基本定理保证我们有以下结论: **如果 R 是有界闭集合, 并且 f 在 R 上连续, 那么 R 内存在着这样的点, f 在其上达到最大, 并且也有这样的点, f 在其上达到最小.**

作为一个例子, 在闭圆盘 $x^2 + y^2 \leq 1$ 上考虑函数 $u = x^2 + y^2$. 这个曲面 S 是旋转抛物面 $z = x^2 + y^2$ 在平面 $z = 1$ 以下的部分. 这里函数 f 在边界圆 $x^2 + y^2 = 1$ 的每一点上都取最大值, 而在原点, f 有一个**严格最小**.

---

1) 也叫做**绝对最大**以区别后面定义的**相对最大**(极大). 这里所用术语与单元函数中一样.

微积分直接运用于求相对极大或极小而非绝对极值，定义域 R 内的一点 (x_0, y_0) 称为相对极大点，如果对 (x_0, y_0) 点的一个充分小的邻域内的所有 R 的点 (x, y) 都有 $f(x_0, y_0) \geq f(x, y)$. 在一个相对极大点上的值 $f(x_0, y_0)$ 不一定是 f 在整个 R 上的最大值，但是如果我们限于考虑充分接近 (x_0, y_0) 的点，它就是 f 的一个最大值. 相对极小可以类似地定义. 每一个绝对极大 (小) 也是一个相对极大 (小), 但是反过来却不成立.

例如，函数 $u = (x^2 + y^2)^3 - 3(x^2 + y^2)$, 以开圆盘

$$x^2 + y^2 < 4$$

为定义域，就没有最大值，是在原点确有一个相对极大值. 圆

$$x^2 + y^2 = 1$$

上的点都是最小点. 这里，曲面 S 是由曲线 $z = x^6 - 3x^2$ 绕 z 轴旋转而产生的.

对于更多个自变量的函数 $u = f(x, y, z, \cdots)$ 的最小与相对极小的定义完全与上述类似.

我们将首先给出在函数 $f(x, y)$ 的定义域 R 的内点 (x_0, y_0) 上取极大或极小的必要条件，我们用术语 "极值" 来概括极大与极小. 现在设 (x_0, y_0) 是函数 $f(x, y)$ 的定义域 R 的一个内点，而且 f 在该点有偏微商 $f_x(x_0, y_0)$, $f_y(x_0, y_0)$. 若要 $f(x, y)$ 在 (x_0, y_0) 点取得一个极值，就必须有

$$f_x(x_0, y_0) = 0, \quad f_y(x_0, y_0) = 0. \tag{67a}$$

条件 (67a) 可以直接由熟知的一元函数的条件推出. 设

$$\phi(x) = f(x, y_0).$$

于是 $\phi(x)$ 对于充分接近 x_0 点的全部 x 都有定义，而且在 x_0 有微商 $\phi'(x_0) = f_x(x_0, y_0)$. 如果对于 R 内充分接近 (x_0, y_0) 点的所

有 (x,y) 点都有 $f(x_0,y_0)\ge f(x,y)$, 那么, 特别地, 对于充分接近 x_0 的全部 x 都有 $\phi(x_0)\ge\phi(x)$. 由此推出 $\phi'(x_0)=0$; 也就是 $f_x(x_0,y_0)=0$. 第二个必要条件

$$f_y(x_0,y_0)=0$$

可以类似地证明.

从几何上来看, $f(x,y)$ 在 (x_0,y_0) 点的两个偏微商等于零的意思就是曲面 $z=f(x,y)$ 在 $(x_0,y_0,f(x_0,y_0))$ 点的切平面平行于 xy 平面. 我们称 (x_0,y_0) 点为函数 $f(x,y)$ 的一个**稳定**点或**临界**点, 如果两个一次偏微商 $f_x(x_0,y_0)$, $f_y(x_0,y_0)$ 都存在而且都等于零. 因此, 可微函数 f 在定义域内部的每一个极值点都是 f 的一个临界点.

同样的结果适用于任意多个自变量的函数 $f(x,y,z,\cdots)$. 这里 $(x_0,y_0,z_0,\cdots)$ 称为 f 的**稳定**点或**临界**点, 如果全部一次偏微商 $f_x, f_y, f_z,\cdots$ 在该点都存在而且满足

$$f_x(x_0,y_0,z_0,\cdots)=0,\quad f_y(x_0,y_0,z_0,\cdots)=0,$$

$$f_z(x_0,y_0,z_0,\cdots)=0,\cdots. \tag{67b}$$

条件的个数与自变量 $x,y,z,\cdots$ 的个数相等. 我们可以把这些条件合并成为一个条件: 对于 $(x,y,z,\cdots)=(x_0,y_0,z_0,\cdots)$ 与一切 $dx,dy,dz,\cdots$ 都有

$$df=f_xdx+f_ydy+f_zdz+\cdots=0.$$

因为方程 (67b) 的个数与未知数 $x_0,y_0,z_0,\cdots$ 的个数相同, 通常可以求得一定个数临界点, 虽然这并非总是这样. 另外, 一个临界点从任何意义上来看都不必是一个极值点.

作为例子, 考虑函数 $u=xy$. (67a) 的两个方程立刻给出点 $(x,y)=(0,0)$ 为唯一的临界点. 但是在 (0,0) 点的每个邻域内, 函数既可以取得正值又可以取得负值, 这决定于 (x,y) 点所在的象

限. 因此, 这个函数在该点没有极值. 表示函数 $u = xy$ 的曲面在几何上是一张双曲抛物面, 这个曲面既无最高点也无最低点, 但是有一个鞍点在原点 (见图 3.1).

我们看到, 一个可微函数的最大点与最小点或者落在函数定义域的边界上, 或者可以在函数的临界点中间找到. 判别一个临界点是最大或最小需要特别的讨论. 在第 385 页上我们将见到保证一个临界点至少是一个极值点的充分条件.

函数 f 的 **最大值** M 是 f 在它的定义域 R 上取得的全部值中的最大者. f 的最大点是使 $f(x, y) = M$ 的那些点. [1] 类似地, f 的 **临界值** 或 **稳定值** 是临界点或稳定点上取得的那些值.

b. 例

1. 函数

$$u = \sqrt{1 - x^2 - y^2} \qquad (x^2 + y^2 < 1)$$

有偏微商

$$u_x = -\frac{x}{\sqrt{1 - x^2 - y^2}},$$

$$u_y = -\frac{y}{\sqrt{1 - x^2 - y^2}},$$

而且这些偏微商在原点都等于零. 这里我们得到了最大值, 因为对于原点附近的所有其他点 (x, y), 平方根下的量 $1 - x^2 - y^2$ 都小于它在原点的值.

2. 我们想作一个三角形, 使得它的三个角的正弦的乘积为最大; 也就是, 我们想求函数

$$f(x, y) = \sin x \sin y \sin(x + y)$$

---

1) 有时候 "最大" 一词笼统地指最大值或最大点.

在区域 $0 \leq x \leq \pi$, $0 \leq y \leq \pi$, $0 \leq x+y \leq \pi$ 上的最大值. 因为 f 在这个区域的内部是正的, 所以它的最大值是正的. 在区域的边界上, 由于规定区域的不等式中至少有一个等号成立, 所以 $f(x,y)=0$, 因此最大值必定在区域内取得.

如果我们令偏微商等于零, 就得到两个方程

$$\cos x \sin y \sin(x+y)+\sin x \sin y \cos(x+y)=0,$$

$$\sin x \cos y \sin(x+y)+\sin x \sin y \cos(x+y)=0.$$

因为 $0<x<\pi$, $0<y<\pi$, $0<x+y<\pi$, 这些方程给出 $\tan x=\tan y$, 或 $x=y$. 如果我们把这个值代入第一个方程, 我们就得到关系式 $\sin 3x=0$; 所以 $(x,y)=\left(\frac{\pi}{3},\frac{\pi}{3}\right)$ 是唯一的稳定点, 因此所求的三角形是等边的.

3. 设 P_1,P_2,P_3 是一个锐角三角形的顶点, 坐标分别为 (x_1,y_1), (x_2,y_2), (x_3,y_3). 如果我们想求第四个点 P, 其坐标为 (x,y), 使得它与 P_1,P_2,P_3 的距离之和为最小. 这个距离之和是 x 与 y 的连续函数, 而且在包含这个三角形的一个大圆内的某一点 P 达到最小值. 这个 P 点不可能落在三角形的一个顶点上, 因为否则其他两个顶点中的任何一个顶点到它的对边的垂线的垂足将有更小的距离之和. 还有, 如果这个圆的圆周离三角形充分远, 那么 P 点也不可能落在圆周上. 设距离 r_i 为

$$r_i=\sqrt{(x-x_i)^2+(y-y_i)^2}$$

我们想使函数

$$f(x,y)=r_1+r_2+r_3$$

取得最小, 这个函数除 P_1,P_2, 与 P_3 点之外处处是可微的, 我们知道这个函数对 x 与对 y 的偏微商在 P 点都必须等于零. 因此, 对 f 进行微分, 我们就得到对于 P 点的条件:

$$\frac{x-x_1}{r_1}+\frac{x-x_2}{r_2}+\frac{x-x_3}{r_3}=0,$$

$$\frac{y-y_1}{r_1}+\frac{y-y_2}{r_2}+\frac{y-y_3}{r_3}=0.$$

根据这些条件，下面这三个平面向量

$$\left(\frac{x-x_1}{r_1},\frac{y-y_1}{r_1}\right),\quad \left(\frac{x-x_2}{r_2},\frac{y-y_2}{r_2}\right),$$

$$\left(\frac{x-x_3}{r_3},\quad \frac{y-y_3}{r_3}\right)$$

的和为 0. 这些向量也都是单位向量. 给定这些向量一个公共的起点 P, 它们的终点就形成一个等边三角形；也就是，每一个向量的方向到另一个向量的方向相差 $2\pi/3$(图 3.27). 由于这三个向量与从 P 到 P_1,P_2,P_3 的三个向量具有一致的方向，这意味着三角形的每一条边都必须在 P 点对着相同的 $2\pi/3$ 角.

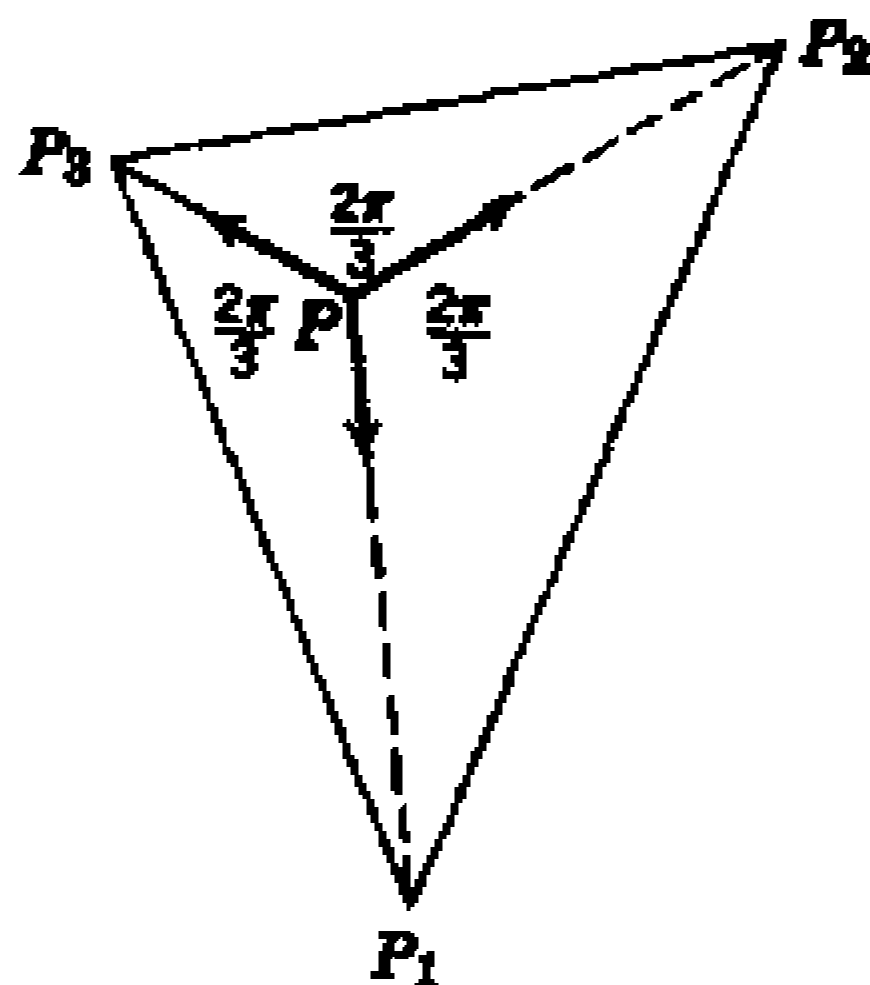


图　3.27

练　习　3.7 b

1. 试求下列函数的稳定点，并指明其类型：

(a) $f(x,y)=y^2(\sin x-x/2)$,

(b) $f(x,y)=\cos(x+y)+\sin(x-y)$,

(c) $f(x,y)=y^x$,

(d) $f(x,y)=x/y$,

(e) $f(x,y)=ye^{-x^2}$.

2. 试求函数

$$(ax^2+by^2)e^{-x^2-y^2}$$

的最大与最小.

3. 试求 $2x^3+(x-y)^2-6y$ 的稳定点.

4. 设长方体的 12 条棱的长度总和为 a; 它的 6 个面的面积总和为 $a^2/25$. 试求: 当长方体的体积与其最短的棱所构成的立方体体积之差为最大时各棱的长度.

5. 试求函数

$$f(x,y,z)=x^2(y-1)^2\left(z+\frac{1}{2}\right)^2$$

的稳定点, 并指明其类型.

6. 按美国现行的邮政条令, 一个边长为 x,y,z $(x\leq y\leq z)$ 的长方形包裹, 只有当 $2(x+y)+z\leq 100$ 时才能邮寄. 试在这个条件下求可邮寄的包裹的最大体积. [提示: 令 $z=100-2(x+y)$.]

7. 试求点 X, 使其与 n 个已知点的距离的平方之和为最小.

c. 带有附加条件的最大与最小

多元函数求最大与最小的问题常常具有各种不同的形式. 例如, 我们可以在给定的曲面 $\phi(x,y,z)=0$ 上求最靠近原点的点. 那么我们就要使函数

$$f(x,y,z)=\sqrt{x^2+y^2+z^2}$$

取得最小. 然而这里量 x,y,z 不再是三个独立的变量, 而是被曲面方程 $\phi(x,y,z)=0$ 这个附加条件所约束. 这种带有附加条件的最大与最小, 事实上并不是根本性的新问题. 例如在我们的例子中, 我们只须把其中一个变量, 譬如说 z, 解成其余两个变量的函数, 问题就归结为求两个自变量 x,y 的函数的稳定值了.

然而，把稳定值的条件表成对称形式，而不偏向于任何一个变量，那就更加方便，同时也更加完善.

一个简单的典型情形就是求函数 $f(x,y)$ 的稳定值，其中两个变量 x,y 并不互相独立，而是被附加条件

$$\phi(x,y)=0$$

所约束. 为了给以几何直观，我们首先假定附加条件是 xy 平面上一条没有奇点的曲线 (见图 3.28)，另外还假定曲线族 $f(x,y)=c=$ 常数覆盖着平面的一部分，如图 3.28. 在曲线族与曲线 $\phi(x,y)=0$

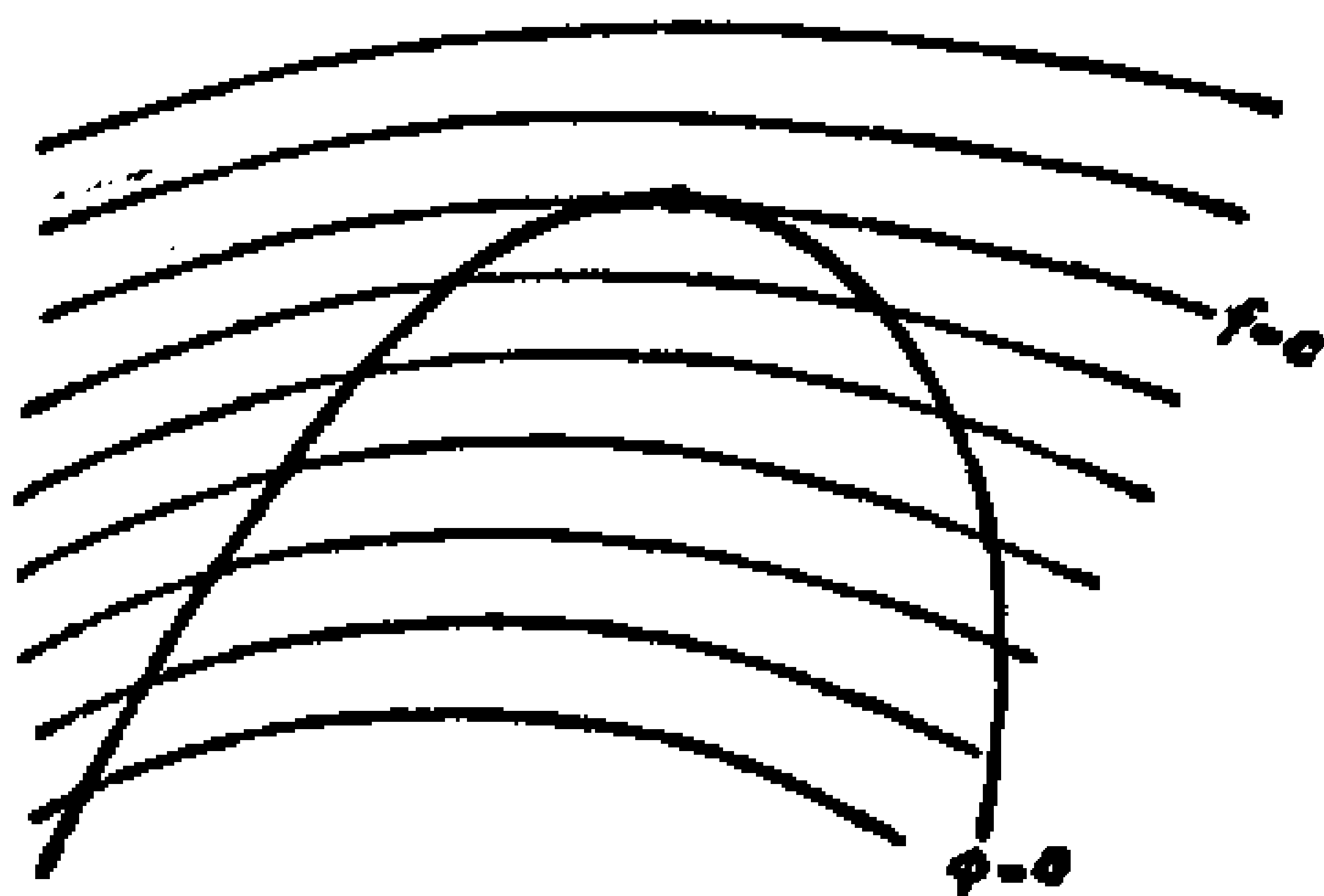


图 3.28　具有附加条件 $\phi=0$ 的 f 的极值

相交的那些曲线中，我们需要找到一条曲线，其对应的常数 c 是最大或最小. 当我们描曲线 $\phi(x,y)=0$ 的时候，我们要穿过曲线族 $f(x,y)=c$；这时，一般说来，c 的变化是单调的；一旦当我们描到 c 改变其单调性的点上，我们可以预期这是一个极值. 从图 3.28，我们看到这发生在曲线族的一条曲线与曲线 $\phi=0$ 相切处. 切点的坐标 $(x,y)=(\xi,\eta)$ 就对应着所要求的 $f(x,y)$ 的极值. 如果两条曲线 $f=$ 常数与 $\phi=0$ 相切，它们就有相同的切线. 因此，在点 $(x,y)=(\xi,\eta)$ 处有比例式

$$f_x:f_y=\phi_x:\phi_y$$

成立；或者，如果我们引进比例常数 λ，就有联立方程

$$\begin{cases} f_x+\lambda\phi_x=0, \\ f_y+\lambda\phi_y=0. \end{cases}$$

这两个方程以及方程

$$\phi(x,y)=0,$$

可以用来决定切点的坐标 (ξ,η) 与比例常数 λ.

上述论证法有可能行不通的时候，例如，当曲线 $\phi(x,y)=0$ 在点 (ξ,η) 有一个奇点，譬如说像图 3.29 中那样的尖点时，它在这点与曲线 $f(x,y)=c$ 相遇，而 c 为最大或最小的可能值. 不过，在这种情形，我们同时有

$$\phi_x(\xi,\eta)=0,\quad \phi_y(\xi,\eta)=0.$$

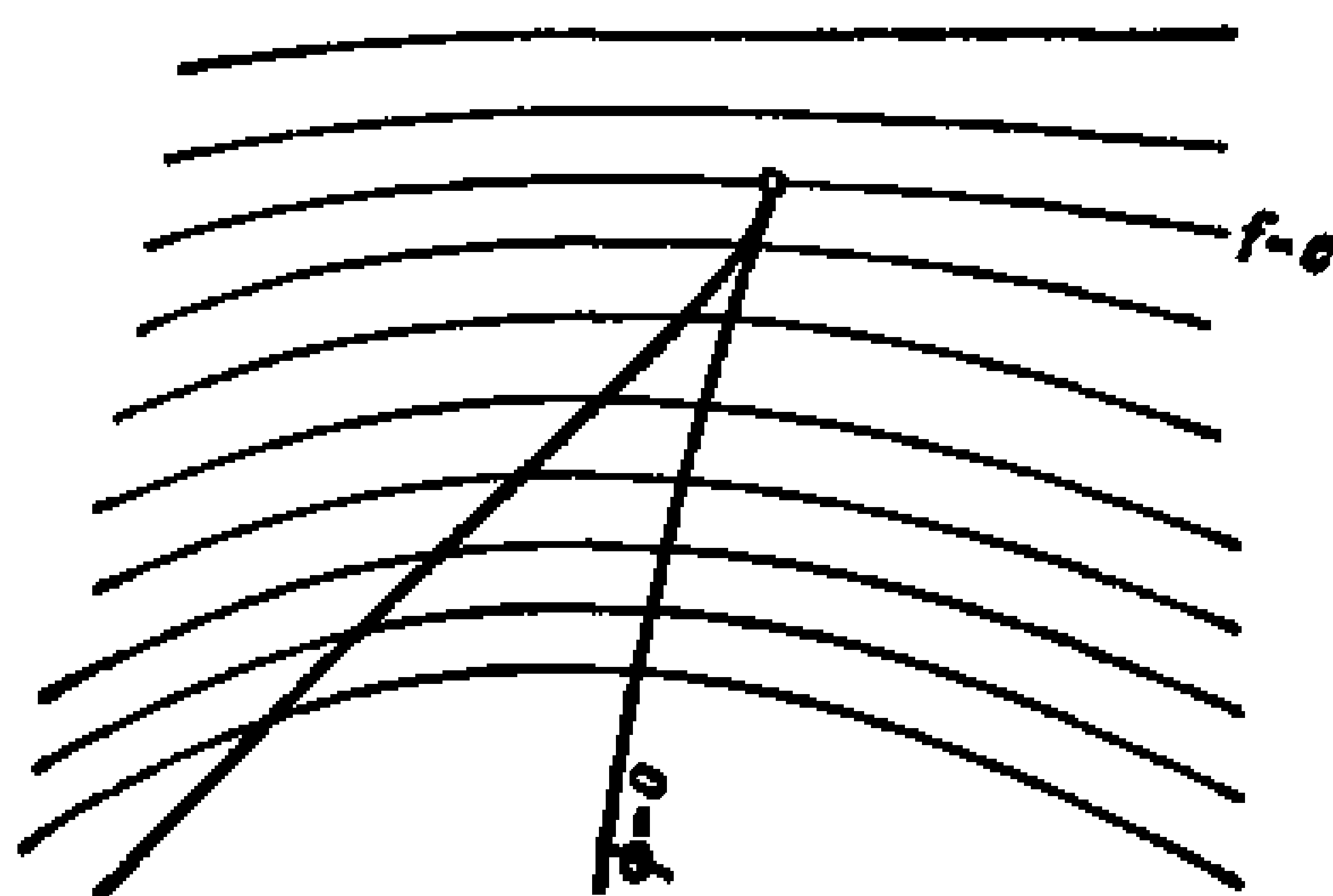


图 3.29　极值取在 $\phi=0$ 的奇点上

我们从直观上得到了下述法则 (其证明将在下一小节中给出):

若要函数 $f(x,y)$ 在附加条件 $\phi(x,y)=0$ 下在点 $(x,y)=(\xi,\eta)$ 处取得极值，那么在 $\phi_x(x,y)$ 与 $\phi_y(x,y)$ 不同时为零的前提下，必须有一个比例常数 λ 存在，使得方程

$$f_x(\xi,\eta)+\lambda\phi_x(\xi,\eta)=0 \text{ 与 } f_y(\xi,\eta)+\lambda\phi_y(\xi,\eta)=0 \tag{67c}$$

以及方程

$$\phi(\xi,\eta)=0 \tag{67d}$$

同时得到满足.

这个准则称为拉格朗日不定乘数法，而因子 λ 称为拉格朗日乘数.

我们注意到这个法则给出的求量 ξ,η 与 λ 的方程的个数等于未知数的个数. 于是我们就把求极值点 (ξ,η) 这个问题变成了带有一个附加未知量 λ 但是具有完全对称性的一般极值问题.

拉格朗日法则通常表述如下:

为了求函数 $f(x,y)$ 在带有附加条件 $\phi(x,y)=0$ 下的极值，我们把函数 $f(x,y)$ 加上 $\phi(x,y)$ 与一个与 (x,y) 无关的未知因子 λ 的乘积，并且写下这些求函数 $F=f+\lambda\phi$ 极值的必要条件.

$$f_x+\lambda\phi_x=0,\quad f_y+\lambda\phi_y=0.$$

结合附加条件 $\phi(x,y)=0$, 就可以用来确定极值点的坐标与比例常数.

作为一个例子，我们来求函数

$$u=xy$$

在以原点为中心的单位圆周上，即具有附加条件

$$x^2+y^2-1=0$$

的极值. 按照我们的法则，把 $xy+\lambda(x^2+y^2-1)$ 对 x, 对 y 求微商，我们求得在稳定点处，这两个方程

$$y+2\lambda x=0,$$

$$x+2\lambda y=0$$

要满足. 此外我们还有的附加条件

$$x^2+y^2-1=0.$$

解出方程，我们得到四个点：

$$\xi=\frac{\sqrt{2}}{2},\quad \eta=\frac{\sqrt{2}}{2};$$
$$\xi=-\frac{\sqrt{2}}{2},\quad \eta=-\frac{\sqrt{2}}{2};$$
$$\xi=\frac{\sqrt{2}}{2},\quad \eta=-\frac{\sqrt{2}}{2};$$
$$\xi=-\frac{\sqrt{2}}{2},\quad \eta=\frac{\sqrt{2}}{2}.$$

前两个点给出函数 $u=xy$ 的最大值 $u=\frac{1}{2}$，而后两个点给出最小值 $u=-\frac{1}{2}$. 至于这前两个点确实给出函数 u 的最大值后两个点确实给出函数的最小值 —— 这个结论的根据则在于这个事实：因为圆周是闭集，而且有界，所以函数在圆周上必定达到最大值与最小值.

练 习 3.7 c

1. 试把 3.7b 节的习题 6 当作在附加条件

$$2(x+y)+z=100$$

下求体积的最大值的问题来求解.

2. 试求函数 $z=x^2y^2$ 在条件 $x+y=1$ 下的最小值.

3. 试求函数 $z=\cos\pi(x+y)$ 在条件 $x^2+y^2=1$ 下的最大值.

4. 试在平面上求一点 X，使其与 n 个已知点的距离的平方和为最小，附加条件为 X 点在一给定的直线上 (与第 3.7b 节的习题 7 作比较).

5. 如果 $C=f(a,b)$ 是函数 $f(x,y)$ 在条件 $\phi(x,y)=C'$ 下的一个真正最大或最小值，求证，一般说来 $C'=\phi(a,b)$ 是函数 $\phi(x,y)$ 在条件 $f(x,y)=C$ 下的一个真正最大或最小值.

d. 最简单情形下不定乘数法的证明

正如我们能预料的，我们作出不定乘数法的分析证明是把它归结为“自由”极值的已知情形而达到的．我们假设在一个极值点处两个偏微商 $\phi_x(\xi,\eta)$ 与 $\phi_y(\xi,\eta)$ 不全为零；为确定起见，我们假设 $\phi_y(\xi,\eta)\neq 0$. 于是，根据隐函数定理（第 239 页），在这个点附近方程 $\phi(x,y)=0$ 唯一地把 y 确定为 x 的一个连续可微的函数 $y=g(x)$. 如果我们把这个表达式代入 $f(x,y)$，函数

$$f(x,g(x))$$

就必定在 $x=\xi$ 点有一个自由极值．因此，方程

$$f'(x)=f_x+f_yg'(x)=0$$

必定在 $x=\xi$ 点成立．此外，隐式确定的函数 $y=g(x)$ 又恒等地满足关系式

$$\phi_x+\phi_yg'(x)=0.$$

如果我们用 $\lambda=-f_y/\phi_y$ 乘这个方程，并把它加到

$$f_x+f_yg'(x)=0$$

上去，我们就得到

$$f_x+\lambda\phi_x=0;$$

而根据 λ 的定义，方程

$$f_y+\lambda\phi_y=0$$

成立．这就证明了不定乘数法．

这个证明揭示了在 (ξ,η) 点导数 ϕ_x 与 ϕ_y 不全为零的假定的重要性；下述例子就说明，如果这两个偏微商都为零，这法则就不成立．

我们要使函数

$$f(x,y)=x^2+y^2$$

在附加条件

$$\phi(x,y)=(x-1)^3-y^2=0$$

下达到最小. 在图 3.30 中, 原点到曲线 $(x-1)^3-y^2=0$ 的最短

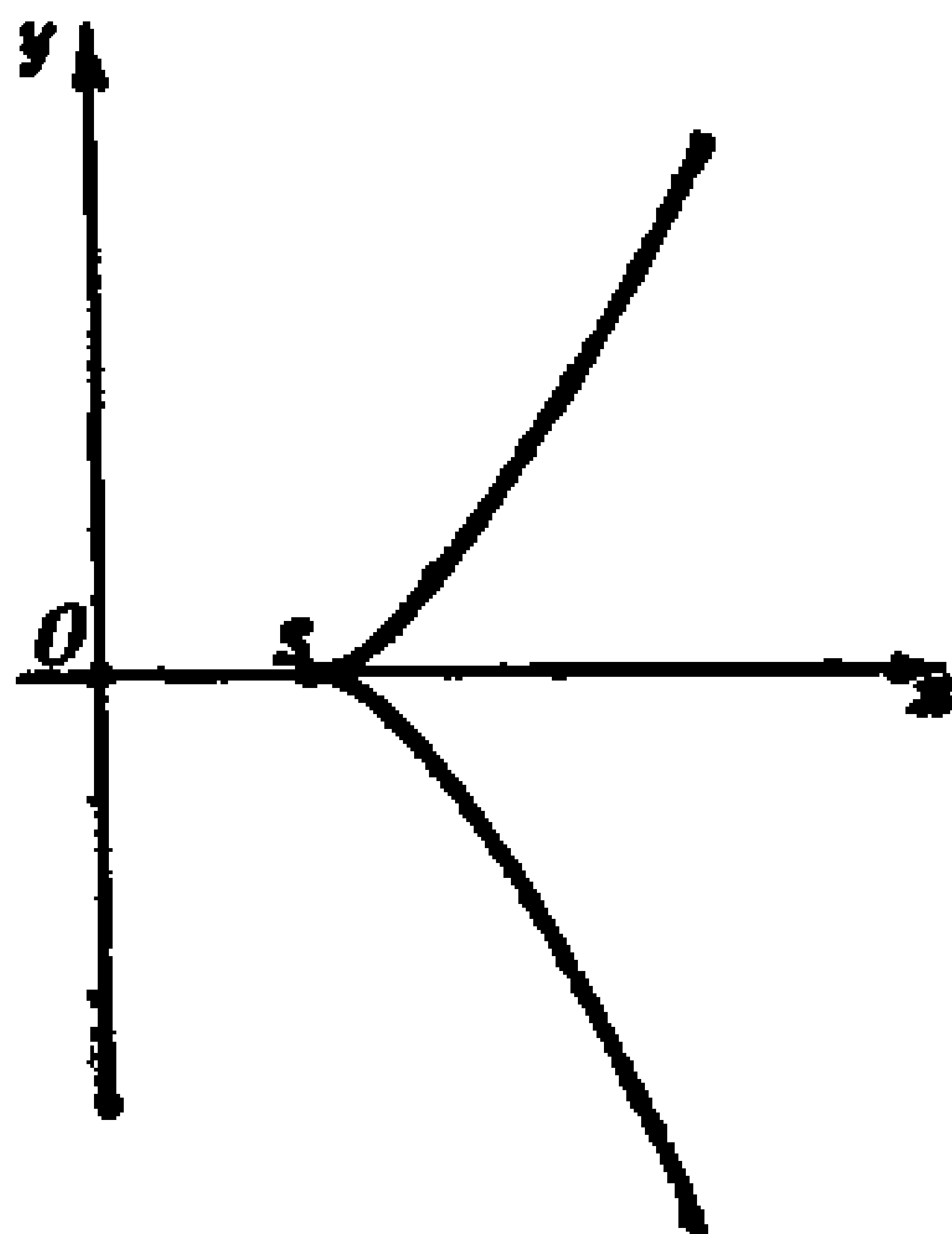


图 3.30 曲线 $(x-1)^3-y^2=0$

距离显然就是连接原点与曲线的尖点 S 的线段 (我们容易证明, 在以原点为中心的单位圆上再没有曲线的别的点), 点 S 的坐标, 即 $x=1$ 与 $y=0$, 满足方程 $\phi(x,y)=0$ 与 $f_y+\lambda\phi_y=0$, 不论 λ 为何值都一样, 但是这里却有

$$f_x+\lambda\phi_x=2x+3\lambda(x-1)^2=2\neq 0.$$

我们可以用略微不同的说法来陈述不定乘数法, 这种说法特别便于推广. 我们曾经看到, 函数 $F(x,y)$ 在某一点的微分为零是函数在该点取得自由极值的一个必要条件. 对于当前的问题, 我们可以类似地作以下的陈述:

若要函数 $f(x,y)$ **在点** (ξ,η) **具有适合附加条件**

$$\phi(x,y)=0$$

的极值，那么在该点处微分 df 必须为零，其中微分 dx 与 dy 看作不是独立的，而是适合由 $\phi=0$ 推得的方程

$$d\phi=\phi_x dx+\phi_y dy=0. \tag{67e}$$

假设在点 (ξ,η) 处微分 dx 与 dy 只要满足方程 $\phi=0$ 的时候就满足方程

$$df=f_x(\xi,\eta)dx+f_y(\xi,\eta)dy=0. \tag{67f}$$

把方程 (67e) 乘以一个数 λ 并加到 (67f)，就得到

$$(f_x+\lambda\phi_x)dx+(f_y+\lambda\phi_y)dy=0.$$

如果我们取 λ 使得

$$f_y+\lambda\phi_y=0 \tag{67g}$$

(这在 $\phi_y\neq 0$ 的假定下是可以做到的)，就立即推出

$$(f_x+\lambda\phi_x)dx=0,$$

而由于 (67e) 中的微分 dx 可以任意选取，例如取它为 1，我们就有

$$f_x+\lambda\phi_x=0. \tag{67h}$$

反之，关系式 (67g,h) 对于任何 λ 成立当然就意味着 $df=0$ 以及 $d\phi=0$.

练　习　3.7 d

1. 对于拉格朗日乘数 λ 与约束条件 $\phi=0$，描述曲面

$$z=f(x,y)+\lambda\phi(x,y)$$

的状态.

e. 不定乘数法的推广

我们可以把不定乘数法推广到多个变量而且也可以推广到多个附加条件. 我们将考虑一个特殊情形, 但这种情形包含了所有本质的要点. 我们来求函数

$$u = f(x, y, z, t) \tag{68a}$$

的极值, 其中四个变量 x, y, z, t 满足两个附加条件

$$\phi(x, y, z, t) = 0, \quad \psi(x, y, z, t) = 0. \tag{68b}$$

我们假设在点 (ξ, η, ζ, τ) 处函数 f 取得一个值, 这个值与所有满足附加条件的邻近点上的值相比是一个极值. 我们要求在点 $P = (\xi, \eta, \zeta, \tau)$ 的附近, 有两个变量, 譬如说是 z 与 t, 可以通过方程 (68b) 表成另外两个变量 (x, y) 的函数. 为要保证能够得到这样的解 $z = g(x, y)$ 与 $t = h(x, y)$, 我们假设在点 P 处雅可比行列式

$$\frac{d(\phi, \psi)}{d(z, t)} = \phi_z \psi_t - \phi_t \psi_z \tag{68c}$$

不等于零 (见 290 页). 我们现在把函数

$$z = g(x, y) \quad 与 \quad t = h(x, y)$$

代入函数 $u = f(x, y, z, t)$, 得到一个依赖于两个自变量 x 与 y 的函数, 并且这个函数在点 $(x, y) = (\xi, \eta)$ 处必须取得一个自由极值; 也就是它的两个偏微商在该点处必须都等于零. 因此这两个方程

$$f_x + f_z \frac{\partial z}{\partial x} + f_t \frac{\partial t}{\partial x} = 0, \tag{69a}$$

$$f_y + f_z \frac{\partial z}{\partial y} + f_t \frac{\partial t}{\partial y} = 0 \tag{69b}$$

必须同时成立. 为了从附加条件中算出在这里出现的四个偏微商 $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial t}{\partial x}, \frac{\partial t}{\partial y}$, 我们可以试图写出这样两对方程

$$\phi_x + \phi_z \frac{\partial z}{\partial x} + \phi_t \frac{\partial t}{\partial x} = 0, \tag{69c}$$

$$\psi_x + \psi_z \frac{\partial z}{\partial x} + \psi_t \frac{\partial t}{\partial x} = 0 \tag{69d}$$

与

$$\phi_y + \phi_z \frac{\partial z}{\partial y} + \phi_t \frac{\partial t}{\partial y} = 0, \tag{69e}$$

$$\psi_y + \psi_z \frac{\partial z}{\partial y} + \psi_t \frac{\partial t}{\partial y} = 0, \tag{69f}$$

并且由这些方程把 $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \frac{\partial t}{\partial x}, \frac{\partial t}{\partial y}$ 作为未知数解出来 $\Big($ 因为雅可比行列式 $\frac{d(\phi,\psi)}{d(z,t)} \neq 0$, 所以是可能的 $\Big)$. 这样问题就解决了.

但是我们不这样做, 我们宁愿为了保持公式的对称性而进行如下的推导. 我们用这样的方法来决定两个数 λ 与 μ, 即在取得极值的点上两个方程

$$f_z + \lambda\phi_z + \mu\psi_z = 0, \tag{70a}$$

$$f_t + \lambda\phi_t + \mu\psi_t = 0 \tag{70b}$$

都满足 (这样来确定乘数 λ 与 μ 确实是可能的, 因为已经假设了雅可比行列式 $\frac{d(\phi,\psi)}{d(z,t)} \neq 0$). 如果我们分别用 λ 与 μ 乘方程 (69c, d), 并且把它们加到方程 (69a) 上, 我们就有

$$f_x + \lambda\phi_x + \mu\psi_x + (f_z + \lambda\phi_z + \mu\psi_z) \\ \times \frac{\partial z}{\partial x} + (f_t + \lambda\phi_t + \mu\phi_t)\frac{\partial t}{\partial x} = 0.$$

因此, 根据确定 λ 与 μ 的 (70a,b), 我们有

$$f_x + \lambda\phi_x + \mu\psi_x = 0.$$

类似地, 如果我们分别用 λ 与 μ 乘方程 (69e,f), 并且把它们加到方程 (69b) 上, 我们就又得到一个方程

$$f_y + \lambda\phi_y + \mu\psi_y = 0.$$

这样我们就得出下面的结果：如果点 (ξ,η,ζ,τ) 是函数 $f(x,y,z,t)$ 在附加条件

$$\phi(x,y,z,t)=0, \tag{71a}$$

$$\psi(x,y,z,t)=0 \tag{71b}$$

下的一个极值点，而且在该点 $\dfrac{d(\phi,\psi)}{d(z,t)}$ 不为零，那么就存在两个数 λ 与 μ，使得在点 (ξ,η,ζ,τ) 处方程

$$f_x+\lambda\phi_x+\mu\psi_x=0, \tag{72a}$$

$$f_y+\lambda\phi_y+\mu\psi_y=0, \tag{72b}$$

$$f_z+\lambda\phi_z+\mu\psi_z=0, \tag{72c}$$

$$f_t+\lambda\phi_t+\mu\psi_t=0, \tag{72d}$$

以及附加条件(71a,b)都满足.

最后这组条件是完全对称的，特别偏重于 x 与 y 两个变量的任何痕迹都不见了；而且，如果我们假设雅可比行列式

$$\frac{\partial(\phi,\psi)}{\partial(x,y)},\frac{\partial(\phi,\psi)}{\partial(x,z)},\cdots,\frac{\partial(\phi,\psi)}{\partial(z,t)}$$

中至少有一个不为零来代替假设 $\dfrac{\partial(\phi,\psi)}{\partial(z,t)}\neq 0$，从而在所讨论的点的邻域内，量 x,y,z,t 中有两个 (不必是 z 与 t) 可以被其他两个所表示，那么我们可以同样得到 (72a,b,c,d). 对于我们的这些方程的这个对称性，我们当然是付了代价的；在未知数 ξ,η,ζ,τ 之外现在又增加了 λ 与 μ. 这样，代替四个未知数，我们现在有六个，由上面的六个方程来确定.

用完全相同的方法，我们可以陈述并证明对于任何个数变量与任何个数附加条件的不定乘数法. 这个普遍法则如下：

如果在一个函数

$$f(x_1, x_2, \cdots, x_n)$$

中，这 n 个变量 $x_1, x_2, \cdots, x_n$ 不是独立的，而是被 m 个条件 $(m < n)$

$$\phi_1(x_1, x_2, \cdots, x_n) = 0,$$

$$\phi_2(x_1, x_2, \cdots, x_n) = 0,$$

$$\cdots\cdots$$

$$\phi_m(x_1, x_2 \cdots, x_n) = 0$$

所约束，我们就引进 m 个乘数 $\lambda_1, \lambda_2, \cdots, \lambda_m$，并且令函数

$$F = f + \lambda_1\phi_1 + \lambda_2\phi_2 + \cdots + \lambda_m\phi_m$$

对于 $x_1, x_2, \cdots, x_n$ 的偏微商(其中 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 都是常数)都为零. 这样得到的方程[1]

$$\frac{\partial F}{\partial x_1} = 0, \frac{\partial F}{\partial x_2} = 0, \cdots, \frac{\partial F}{\partial x_n} = 0,$$

连同 m 个附加条件

$$\phi_1 = 0,\ \phi_2 = 0, \cdots, \phi_m = 0$$

是对于 $m+n$ 个未知数 $x_1, x_2, \cdots, x_n; \lambda_1, \lambda_2, \cdots, \lambda_m$ 的一个 $m+n$ 个方程的方程组. 这些方程都必须在 f 的任何极值点上得到满足，除非在该点处这 m 个函数 $\phi_1, \phi_2, \cdots, \phi_m$ 对于变量

$$x_1, x_2, \cdots, x_n$$

1) 这些方程与对于辅助函数 F 的自由极值的方程完全相同.

中任何 m 个变量的雅可比行列式统统都等于零.

我们注意到，这个法则给了我们一个确定极值点的完美形式的方法；然而它仅仅确立了一个必要条件. 还需要讨论的是，我们用不定乘数法求出的点是否确实对应于函数的最大或最小. 对于这个问题，我们将不去进行讨论；因为它将使我们离题太远. 像在自由极值的情形一样，我们运用不定乘数法时，我们通常已经预先知道在 f 的定义域内部有一个极值点存在. 如果这个方法确定出唯一的点并且例外情形 (所有雅可比行列式为零) 在所讨论区域内的任何地方都不发生，那么我们就可以确信我们确实已经找到了极值点.

练　习　3.7 e

1. 试对函数 $u = f(x,y,z)$ 在条件 $\phi(x,y,z) = 0$ 下的最小值问题给以几何解释.

2. 试给出下述形式的问题的一个例子： $f(x,y,z)$ 适合条件 $\phi(x,y) = 0, \psi(y,z) = 0$ 的极值. 对此给以几何解释.

f. 例

1. 作为第一个例子，我们试图求函数 $f = x^2y^2z^2$ 在附加条件 $x^2 + y^2 + z^2 = c^2$ 下的最大值. 在球面 $x^2 + y^2 + z^2 = c^2$ 上，因为它是一个有界的闭集，所以函数必定达到一个最大值. 按照法则，我们作出表达式

$$F = x^2y^2z^2 + \lambda(x^2 + y^2 + z^2 - c^2),$$

并通过微分法得到

$$2xy^2z^2 + 2\lambda x = 0,$$

$$2x^2yz^2 + 2\lambda y = 0,$$

$$2x^2y^2z + 2\lambda z = 0.$$

带有 $x=0, y=0$ 或 $z=0$ 的解可以除外，因为在这些点处函数 f 取得它的最小值 0. 方程的其他解是 $x^2=y^2=z^2$, $\lambda=-x^4$. 利用附加条件，我们得到所要求的坐标

$$x=\pm\frac{c}{\sqrt{3}}, y=\pm\frac{c}{\sqrt{3}}, z=\pm\frac{c}{\sqrt{3}}.$$

在所有这些点处，函数达到相同的值 $c^6/27$, 这就是最大值. 因此，任何三个数都满足关系式

$$\sqrt[3]{x^2y^2z^2}\le\frac{c^2}{3}=\frac{x^2+y^2+z^2}{3},$$

这个关系表明，**三个非负数** x^2, y^2, z^2 **的几何平均值决不会大于它们的算术平均值**.

可以类似地对任意个数的正数作出证明：**几何平均值决不会超过算术平均值**[1].

2. 作为第二个例子，我们将寻求已知周长为 $2s$ 的三角形 (它的边为 x, y, z) 的最大面积. 按熟知的海伦公式，面积的平方由公式

$$f(x,y,z)=s(s-x)(s-y)(s-z)$$

给出. 因此我们来求这个函数在附加条件

$$\phi=x+y+z-2s=0$$

下的最大值，其中 x, y, z 被不等式

$$x\ge0,\ y\ge0,\ z\ge0,\ x+y\ge z,\ x+z\ge y,\ y+z\ge x$$

所限制. 在这个闭区域的边界上 (即在这些不等式中有一个变成等式时), 我们总有 $f=0$. 所以 f 的最大值出现在区域内部并且是一个极大值. 我们构成函数

$$F(x,y,z)=s(s-x)(s-y)(s-z)+\lambda(x+y+z-2s),$$

1) 其他的证法见第一卷第 121 页问题 13 或第 358 页问题 11.

并通过微商得到三个条件

$$-s(s-y)(s-z)+\lambda=0,\ -s(s-x)(s-z)+\lambda=0,$$

$$-s(s-x)(s-y)+\lambda=0.$$

令这三个式子相等，化简就得到 $x=y=z=2s/3$; 就是说，所求解是一个等边三角形.

3. 其次我们来证明不等式

$$uv \leq \frac{1}{\alpha}u^{\alpha}+\frac{1}{\beta}v^{\beta}, \tag{73a}$$

其中 $u\geq 0$, $v\geq 0$, $\alpha\geq 0$, $\beta\geq 0$, 并且

$$\frac{1}{\alpha}+\frac{1}{\beta}=1.$$

这个不等式当 u 或 v 为零时无疑是有效的. 因此我们可只限于考虑 u 与 v 都不为零的值，即 $uv\neq 0$. 如果对于一对数 u,v 不等式成立，那么对于所有的数 $ut^{\frac{1}{\alpha}}, vt^{\frac{1}{\beta}}$(其中 t 是一个任意正数) 不等式也是成立的. 因此我们只需要考虑 u,v 的适合 $uv=1$ 的那些值. 因此我们要证明，对于所有正数 u,v, 使得 $uv=1$ 的，都有不等式

$$\frac{1}{\alpha}u^{\alpha}+\frac{1}{\beta}u^{\beta}\geq 1$$

成立.

为此，我们求解函数

$$\frac{1}{\alpha}u^{\alpha}+\frac{1}{\beta}u^{\beta}$$

适合附加条件 $uv=1$ 的最小值问题. 显然这个最小值是存在的，而且在这样一点 (u,v) 上出现，其中 $u\neq 0$, $v\neq 0$. 所以存在这样一个乘数 λ

$$u^{\alpha-1}-\lambda v=0 \quad 与 \quad v^{\beta-1}-\lambda u=0.$$

分别用 u 与 v 乘这些方程，立刻得到 $u^\alpha = \lambda$, $v^\beta = \lambda$. 注意到 $uv = 1$, 即可见上面的结果隐含着 $u = v = 1$. 所以

$$\frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta$$

的最小值就是 $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. 这就是说，不等式

$$\frac{1}{\alpha}u^\alpha + \frac{1}{\beta}v^\beta \geq 1$$

当 $uv = 1$ 时已被证明.

如果在不等式 (73a) 中，我们把 u 与 v 分别换成

$$u = u_i\Big/\Big(\sum_{i=1}^n u_i^\alpha\Big)^{1/\alpha} \quad 与 \quad v = v_i\Big/\Big(\sum_{i=1}^n v_i^\beta\Big)^{1/\beta},$$

其中 $u_1, u_2, \cdots, u_n$, $v_1, v_2, \cdots, v_n$ 是任意的非负数，而且至少有一个 u 与至少有一个 v 不是零，然后对 $i = 1, 2, \cdots, n$ 求和，我们就得到赫尔德尔不等式

$$\sum_{i=1}^n u_i v_i \leq \Big(\sum_{i=1}^n u_i^\alpha\Big)^{1/\alpha}\Big(\sum_{i=1}^n v_i^\beta\Big)^{1/\beta}. \tag{73b}$$

这个不等式对于任意 $2n$ 个数 u_i, v_i 都是成立的，只要

$$u_i \geq 0, \quad v_i \geq 0 \qquad (i = 1, 2, \cdots, n)$$

并且 u 不全为零、v 也不全为零，与指数

$$\alpha > 0,\ \beta > 0,\ \frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

柯西 - 施瓦茨不等式是赫尔德尔不等式当 $\alpha = \beta = 2$ 时的特殊情形.

4. 最后我们在闭曲面

$$\phi(x, y, z) = 0$$

上寻找这样的点，使得它与固定点 (ξ,η,ζ) 的距离为最小. 如果距离最小，那么它的平方也最小. 因此我们考虑函数

$$F(x,y,z)=(x-\xi)^2+(y-\eta)^2+(z-\zeta)^2+\lambda\phi(x,y,z).$$

微分后给出条件

$$2(x-\xi)+\lambda\phi_x=0,\ 2(y-\eta)+\lambda\phi_y=0,$$

$$2(z-\zeta)+\lambda\phi_z=0,$$

或者表成另一形式，

$$\frac{x-\xi}{\phi_x}=\frac{y-\eta}{\phi_y}=\frac{z-\zeta}{\phi_z}.$$

这些方程说明：固定点 (ξ,η,ζ) 是在取得极距的点 (x,y,z) 处的曲面法线上. 所以，为了沿一最短路径从一点走向一个 (可微的) 曲面，我们必须沿着曲面的法线方向走. 当然进一步的讨论是要求判定我们求得的是最大还是最小或什么都不是. 作为例子，可考虑球面内的一个点. 取得极距的点是在通过该点的直径的两端，到其中一端的距离为最小，而到另一端的距离为最大.

练　习　3.7 f

1. 试求平面 $Ax+By+Cz=D$ 与点 (a,b,c) 之间的最短距离.

2. 试求椭圆 $\frac{x^2}{4}+y^2=1$ 上一点到直线 $x+y-4=0$ 的最长与最短距离.

3. 试证表达式

$$\frac{ax^2+2bxy+cy^2}{ex^2+2fxy+gy^2}\quad(eg-f^2>0)$$

的最大值等于对 λ 的方程

$$(ac-b^2)-\lambda(ag-2bf+ec)+\lambda^2(eg-f^2)=0$$

中的较大的根.

4. 计算下列表达式的最大值:
(a) $\dfrac{x^2+6xy+3y^2}{x^2-xy+y^2}$, (b) $\dfrac{x^4+2x^3y}{x^4+y^4}$.

5. 试求 a,b 的值, 使得包含圆 $(x-1)^2+y^2=1$ 在其内部的椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 有最小的面积.

6. 球面 $x^2+y^2+z^2=1$ 上的哪一点与点 (1,2,3) 的距离为最大?

7. 试在椭球面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1$ 上求一点 (x,y,z) 使得
(a) $A+B+C$, (b) $\sqrt{A^2+B^2+C^2}$
是一个最小值, 其中 A,B,C 表示在 (x,y,z) 点 (其中 $x>0$, $y>0$, $z>0$) 的切平面与坐标轴构成的截距.

8. 试求包含在椭球面 $x^2/a^2+y^2/b^2+z^2/c^2=1$ 内体积最大的长方体.

9. 试求包含在椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 内周长最大的矩形.

10. 试在椭圆 $5x^2-6xy+5y^2=4$ 上求一点, 使得原点到该点的切线的距离为最大.

11. 试证椭球

$$ax^2+by^2+cz^2+2dxy+2exz+2fyz=1$$

长轴的长度 l 是方程

$$\begin{vmatrix} a-\dfrac{1}{l^2} & d & e \\ d & b-\dfrac{1}{l^2} & f \\ e & f & c-\dfrac{1}{l^2} \end{vmatrix}=0$$

的最大实根.

12. (a) 试求 $x^ay^bz^c$ 适合条件 $x^k+y^k+z^k=1$ 的最大值, 其中 a,b,c,k 是正的常数, x,y,z 非负.

(b) 由 (a) 的结果，对于任何六个正数，试证不等式

$$\left(\frac{u}{a}\right)^a\left(\frac{v}{b}\right)^b\left(\frac{w}{c}\right)^c \leq \left(\frac{u+v+w}{a+b+c}\right)^{a+b+c}$$

13. 设 $P_1P_2P_3P_4$ 是一个凸四边形，试求一点 O, 使得它到 P_1, P_2, P_3, P_4 的距离之和为最小.

14. 试求具有最大面积的四边形，已知它的四条边的长度为 a, b, c, d.

附　　录

A.1 极值的充分条件

在上一章最大与最小的理论中，我们满足于寻求出现一个极值的必要条件. 在许多实际情形中，这样求得的稳定点的类型是可以通过问题所具有的特定性质来决定的，这就使得我们能够判定它是一个最大或是一个最小. 然而有一个关于出现极值的普遍的**充分**条件仍然是重要的. 这样的判别准则，将在这里对于两个自变量的典型情形展开讨论.

我们考虑一点 (x_0, y_0)，它是函数的一个稳定点，也就是这样一个点，函数的两个一次偏微商在该处为零. 若要在这一点出现极值，必须且只须表达式

$$f(x_0+h, y_0+k) - f(x_0, y_0)$$

对于所有绝对值充分小的 h 与 k 具有相同的符号. 如果我们用泰勒定理把这个表达式展开到三次余项，并且利用方程 $f_x(x_0, y_0) = 0$ 与 $f_y(x_0, y_0) = 0$，我们就得到

$$f(x_0+h, y_0+k) - f(x_0, y_0)$$

$$=\frac{1}{2}(h^2f_{xx}+2hkf_{xy}+k^2f_{yy})+\varepsilon\rho^2,$$

其中 $\rho^2=h^2+k^2$, 而 ε 当 ρ 趋于零时趋于零.

这个式子说明, 在点 (x_0,y_0) 的一个充分小的邻域内, 函数差值 $f(x_0+h,y_0+k)-f(x_0,y_0)$ 的大小主要是由表达式

$$Q(h,k)=ah^2+2bhk+ck^2$$

决定的, 其中为了简便, 我们令

$$a=f_{xx}(x_0,y_0),\ b=f_{xy}(x_0,y_0),\ c=f_{yy}(x_0,y_0).$$

为了研究极值问题, 我们必须考察这个 h 与 k 的二次齐次式或二次型 Q. 我们假定系数 a,b,c 不全为零. 在它们全为零的例外情形, 我们必须把泰勒级数展开到更高次项, 我们将不讨论这个情形.

二次型 Q 有三种不同的可能情形:

1. 二次型是 **定** 型. 这就是当 h 与 k 取所有的值, Q 的值都有相同的符号, 而且只在 $h=0$ 与 $k=0$ 时其值为零. 我们按它的符号是正的或负的而分别称它为 **正定** 或 **负定**. 例如, 表达式 h^2+k^2(即当 $a=c=1,b=0$ 时我们所得到的二次型) 是正定的; 但表达式 $-h^2+2hk-2k^2=-(h-k)^2-k^2$ 是负定的.

2. 二次型是 **不定型**. 就是它可以取不同符号的值; 例如 $Q=2hk$, 对于 $h=1,\ k=1$ 它取值为 2; 而对于 $h=-1,\ k=1$ 它取值为 -2.

3. 第三种可能是二次型 Q 对于不同于 $h=0,\ k=0$ 的某些 h,k 的值取值为零, 但是在其他地方取具有同一符号的值, 例如型 $(h+k)^2$, 它对于所有满足 $h=-k$ 的 h,k 的值集合都为零. 这样的二次型称为 **半定型**.

若要二次型 $Q=ah^2+2bhk+ck^2$ 是定型, 必须且只须它的判别式 $ac-b^2$ 满足条件

$$ac-b^2>0;$$

如果 $a > 0$(因此也有 $c > 0$), 它就是正定的; 否则它是负定的.

要使二次型 Q 是不定型, 必要且充分的条件是

$$ac - b^2 < 0,$$

而半定型情形的特征则由方程 [1)]

$$ac - b^2 = 0$$

来刻画.

现在我们要证明下面的陈述. 如果二次型 $Q(h, k)$ 是正定的. 则对于 $h = 0, k = 0$ 所达到的稳定值是一个**极小值**(一个严格的极小). 如果二次型是负定的, 稳定值就是一个极大值. 如果二次型是不定的, 我们就既没得到极大, 也没得到极小, 这个点是一个**鞍点**. 因此二次型 Q 的"定"的特性是一个关于极值的充分条件, 而 Q 的"不定"的特性排除了极值的可能性. 我们将不考虑半定的情况, 因为这种情况需要麻烦的讨论.

为了证明第一个陈述, 我们注意, 如果 Q 是一个正定型, 就有

---

1) 这些条件容易推得如下: 在 $a = c = 0$ 的情形, 必有 $b \neq 0$, 所以二次型是不定的, 所以判别准则对这种情形是成立的; 另外, 当 $a \neq 0$ 时, 我们可以把二次型写成

$$ah^2 + 2bhk + ck^2 = a\left[\left(h + \frac{b}{a}k\right)^2 + \frac{ca - b^2}{a^2}k^2\right].$$

如果 $ca - b^2 > 0$, 这个二次型显然是定型, 并且与 a 具有同样的符号. 如果 $ca - b^2 = 0$, 这个型就是半定型, 它对于满足方程

$$\frac{h}{k} = -\frac{b}{a}$$

的所有 h, k 值取值为零; 但是对于一切其他的 h, k 值, 它取值具有同一符号. 如果 $ca - b^2 < 0$, 它就是一个不定型, 当 $k = 0$ 与当 $h + \left(\frac{b}{a}\right)k = 0$ 时取不同符号的值.

一个与 h, k 无关的正数 m, 使得 [1]

$$Q \geq 2m(h^2+k^2) = 2m\rho^2,$$

因此

$$\begin{aligned} & f(x_0+h, y_0+k) - f(x_0, y_0) \\ & \quad = \frac{1}{2}Q(h,k) + \varepsilon\rho^2 \geq (m+\varepsilon)\rho^2. \end{aligned}$$

如果我们现在取 ρ 如此之小, 使得数 ε 的绝对值小于 $(1/2)m$, 我们显然就有

$$f(x_0+h, y_0+k) - f(x_0, y_0) \geq \frac{m}{2}\rho^2 > 0.$$

这样, 对于点 (x_0, y_0) 的这个邻域, 函数的值 (当然除去 (x_0, y_0) 点本身) 就处处大于 $f(x_0, y_0)$. 同样地, 当二次型是负定时, 这个点就是极大点.

最后, 如果二次型是不定型, 那么就有一对值 (h_1, k_1) 使 Q 是负的; 而又有另外一对值 (h_2, k_2) 使 Q 是正的. 我们因此可以找到一个正数 m 使得

$$Q(h_1, k_1) < -2m\rho_1^2,$$

$$Q(h_2, k_2) > 2m\rho_2^2.$$

如果我们令 $h = th_1$, $k = tk_1$, $\rho^2 = h^2 + k^2$, $(t \neq 0)$—— 也就是我们考虑连结 (x_0, y_0) 与 (x_0+h_1, y_0+k_1) 的线段上的一个点 (x_0+h, y_0+k) —— 那么 $Q(h,k) = t^2 Q(h_1, k_1)$, $\rho^2 = t^2\rho_1^2$, 因此我们就有

$$Q(h,k) < -2m\rho^2.$$

---

1) 为了说明这一点, 我们把商式 $Q(h,k)/(h^2+k^2)$ 看作是两个量

$$u = h/\sqrt{h^2+k^2} \text{ 与 } v = k/\sqrt{h^2+k^2}$$

的函数. 于是 $u^2+v^2=1$, 并且商式是 u 与 v 的连续函数, 因此必定在圆 $u^2+v^2=1$ 上有一个最小值 $2m$. 这个值 m 显然满足我们的条件; 因为在圆上 u 与 v 不会同时等于零, 所以 m 决不等于零.

这样，只要选取一个充分小的 t(以及对应的 ρ). 我们就可以使表达式 $f(x_0+h,y_0+k)-f(x_0,y_0)$ 是负的. 我们只须选取 t 如此之小，使得对于 $h=th_1$, $k=tk_1$ 量 ε 的绝对值小于 $\frac{1}{2}m$. 对于 h,k 的这样一组值，我们有

$$f(x_0+h,y_0+k)-f(x_0,y_0)<-m\rho^2/2,$$

因而函数值 $f(x_0+h,y_0+k)$ 小于稳定值 $f(x_0,y_0)$. 同样地, 对于 $h=th_2$, $k=tk_2$ 这一组值运用相应的步骤，我们推知，在 (x_0,y_0) 点的一个任意小的邻域内都存在着这样的点, 使函数的值大于 $f(x_0,y_0)$. 我们既没得到极大也没得到极小，因而我们把它称为鞍点.

如果在稳定点上 $a=b=c=0$, 从而二次型恒等于零，因而在半定情形，上面的讨论不适用. 若要对于这些情形获得充分条件，那就要陷入麻烦的讨论.

这样，我们就有了下述关于判定极大与极小的法则:

若在一个点 (x_0,y_0) **处, 偏微商**

$$f_x(x_0,y_0)=0,\qquad f_y(x_0,y_0)=0$$

并且不等式

$$f_{xx}f_{yy}-f_{xy}^2>0$$

成立, 则函数 f **有一个极值. 如果** $f_{xx}<0$(**因此也有** $f_{yy}<0$)**这就是一个极大, 如果** $f_{xx}>0$ **这就是一个极小. 另一方面, 如果**

$$f_{xx}f_{yy}-f_{xy}^2<0,$$

这个稳定值就既不是极大也不是极小, 情形

$$f_{xx}f_{yy}-f_{xy}^2=0$$

留而未决.

这些条件有一个简单的几何解释. 必要条件 $f_x = f_y = 0$ 说明曲面 $z = f(x, y)$ 的切平面是水平的. 如果我们确实有一个极值, 那么在所讨论的这个点的邻域内切平面不会与曲面交叉. 在鞍点的情形中, 与上面相反, 切平面与曲面交叉成一条曲线, 该曲线在这点有若干个分支. 这种事情在 A.3 节奇点的讨论之后将会更加清楚.

作为一个例子, 我们求函数

$$f(x, y) = x^2 + xy + y^2 + ax + by$$

的极值. 如果我们令一次微商等于零, 就得到方程

$$2x + y + a = 0, \quad x + 2y + b = 0,$$

它的解是 $x = \frac{1}{3}(b - 2a)$, $y = \frac{1}{3}(a - 2b)$. 表达式

$$f_{xx}f_{yy} - f_{xy}^2 = 3$$

是正的, 因为 $f_{xx} = 2$. 所以函数在所讨论的点处有一个最小值.

函数

$$f(x, y) = (y - x^2)^2 + x^5$$

有一个稳定点是原点. 这里表达式 $f_{xx}f_{yy} - f_{xy}^2 = 0$, 因而我们的判别法无效. 但是我们容易看出, 这个函数在这里没有极值, 因为在原点的邻域内这个函数值既有正的, 也有负的.

另一方面, 函数

$$f(x, y) = (x - y)^4 + (y - 1)^4$$

有一个最小在点 $(x, y) = (1, 1)$, 虽然在这里 $f_{xx}f_{yy} - f_{xy}^2 = 0$. 因为

$$f(1 + h, 1 + k) - f(1, 1) = (h - k)^4 + k^4,$$

而这个量当 $\rho \neq 0$ 时是正的.

练 习 A.1

1. 试求下列函数的极值并定出其特性:

(a) $f(x,y)=x^2-3xy+y^2$,

(b) $f(x,y)=\cos(x+y)+\sin(x-y)-x^2$,

(c) $f(x,y)=x\cosh y-y^2$.

2. 设 $\phi(a)=k\neq 0$, $\phi'(a)\neq 0$, 而且 x,y,z 满足关系式 $\phi(x)\phi(y)\phi(z)=k^3$, 求证函数 $f(x)+f(y)+f(z)$ 当 $x=y=z=a$ 时有一个最小值, 只要有条件

$$f'(a)\left(\frac{\phi''(a)}{\phi'(a)}-\frac{\phi'(a)}{\phi(a)}\right)>f''(a).$$

3. 设 $P_1P_2P_3$ 是一个平面三角形, 并且三个角都小于 $120°$. 根据第 358 页的判别准则或后面习题 6 求证, 如果在 $P_1P_2P_3$ 内部的一点 P 上, 有 $\angle P_2PP_3=\angle P_3PP_1=\angle P_1PP_2=120°$, 那么和数 $PP_1+PP_2+PP_3$ 是一个真正的最小值 (参看第 361 页例 3).

4. 如果在习题 3 中三角形的 $\angle P_2P_1P_3$ 大于或等于 $120°$, 试问和数 $PP_1+PP_2+PP_3$ 将在何处取得最小值?

5. (a) 试证: 如果所有的字母都表示正的量, 那么和式

$$lx+my+nz$$

在条件

$$x^p+y^p+z^p=c^p$$

下的稳定值是 $c(l^q+m^q+n^q)^{1/q}$, 其中 $q=\dfrac{p}{p-1}$.

(b) 试说明最大值与最小值必定对应于 $p\gtrless 1$.

6. 试把 A.1 节中的研究结果推广到 n 个变量的函数, 证明以下结果: 设函数 $f(x_1,x_2,\cdots,x_m)$ 在其一个稳定点 $x_1=x_1^0,x_2=x_2^0,\cdots,x_n=x_n^0$ (即在该点上有 $f_{x_1}=f_{x_2}=\cdots=f_{x_n}=0$) 的邻域内三次连续可微. 考虑 f 在点 x^0 的二阶全微分

$$d^2f^0=\sum_{i,k=1}^{n}f^0_{x_ix_k}dx_idx_k;$$

这是变量 $dx_1, dx_2, \cdots, dx_n$ 的一个二次型. 如果这个二次型是非退化的, 也就是如果

$$D = \begin{vmatrix} f^0_{x_1x_1} & \cdots & f^0_{x_1x_n} \\ \vdots & & \vdots \\ f^0_{x_nx_1} & \cdots & f^0_{x_nx_n} \end{vmatrix} \neq 0,$$

那么 d^2f^0 可以是 (1) 正定型, (2) 负定型或 (3) 不定型. 试证: 这些可能情形分别对应于 f 在 x^0 点的下述性质: (1) f 有一个最小值; (2) f 有一个最大值, (3) f 既不取最小也不取最大.

7. 考察 $f = f(x_1, x_2, \cdots, x_n)$ 的稳定点, 其中变量满足关系式

$$\phi_1(x_1, x_2, \cdots, x_n) = 0, \cdots, \phi_m(x_1, x_2, \cdots, x_n) = 0 \ (m < n), \quad (1)$$

我们可以假定已经求得变量与乘数 λ_μ 的数值使得

$$F = f + \lambda_1\phi_1 + \cdots + \lambda_m\phi_m$$

满足方程

$$\frac{\partial F}{\partial x_1} = 0, \cdots, \frac{\partial F}{\partial x_n} = 0, \quad (2)$$

并且 $\phi_1, \cdots, \phi_m$ 对变量 $x_1, \cdots, x_m$ 的雅可比行列式不为零. 为了应用习题 6 的判别准则, 我们可以如下进行: 把 $x_{m+1}, \cdots, x_n$ 当作自变量, 对 (1) 进行微分, 我们可以得到作为 $x_{m+1}, \cdots, x_n$ 的函数的 $x_1, \cdots, x_m$ 的一阶与二阶微分, 并把这些值代入

$$d^2f = \sum_{i,k=1}^{n} f_{x_ix_k}dx_idx_k + f_{x_1}d^2x_1 + \cdots + f_{x_m}d^2x_m. \quad (3)$$

试证下述第二法 (则这个法则不计算二次微分 $d^2x_1, \cdots, d^2x_m$): 把 $x_1, \cdots, x_n$ 当作独立变量, 考虑

$$d^2F = \Sigma F_{x_ix_k}dx_idx_k = d^2f + \lambda_1 d^2\phi_1 + \cdots + \lambda_m d^2\phi_m;$$

从方程

$$d\phi_\mu = \phi_{\mu x_1}dx_1 + \cdots + \phi_{\mu x_n}dx_n = 0 \ (\mu = 1, 2, \cdots, m)$$

算出 $dx_1, \cdots, dx_m$ 并把它们代入 d^2F, 这就得到变量 $dx_{m+1}, \cdots, dx_n$ 的一个二次型 δ^2F. 如果这个二次型是非退化的，那么 f 有一个最小值、或一个最大值、或没有最大最小值，必然分别对应于 δ^2F 是正定型、或负定型、或不定型.

8. 在求函数 $f = x_1x_2\cdots x_n$ 在条件 $\phi = x_1 + \cdots + x_n - a = 0$ $(a > 0)$ 下的最大值问题中，不定乘数法给出 f 在点

$$x_1 = x_2 = \cdots = x_n = \frac{a}{n}$$

的一个稳定值. 试用习题 7 中的法则代替关于绝对最大值的考虑来说明 f 在这个点处有一个最大值.

9. 试用习题 7 中的判别准则证明：在周长为常数的所有三角形中等边三角形面积最大 (见第 375 页).

A.2 临界点的个数与向量场的指数

一个定义在有界闭集 R 上的连续函数 $f(x,y)$, 根据我们的基本定理，无疑在 R 内有一个最大点和一个最小点. 如果一个最大或最小点 (x_0, y_0) 是 R 的一个内点，并且 f 在 (x_0, y_0) 是可微的，那么 (x_0, y_0) 是 f 的一个临界点. 在某些情形中，这个结果使我们推知至少存在 f 的一个临界点. 例如，如果集合 R 由一个开的有界集 S 以及它的边界 B 所组成，并且如果 f 在 B 上是常数，又在 S 内可微，那么 f 至少有一个临界点在 S 内. 这正好是罗尔定理 (见第一卷第 196 页) 在多元函数中的一个扩充，而且用同样的方法证明：函数 f 有最大与最小点，如果这些点都在边界 B 上，因为 f 在边界上是常数，所以 f 的最大值与最小值相同，从而 f 在 S 内是常数，也就是每一个 S 内的点都是临界点. 因此，至少有一个 f 的临界点在 S 内.

在一元函数的情形，有较多的关于某种类型的临界点的个数的特殊论断. 在一个区间内，如果一个函数的极大与极小 交替 出现，则全部极大的个数与全部极小的个数就至多相差 1. 这对于定义在平面集合 S 内的二元函数是不对的. 但是无论如何存在着一个 (直观上不明显) R 内全部极值点与鞍点同 f 在 R 的边界上的值之间的关系. 为了确切陈述这个关系，我们需要考虑 f 的 梯度场，并且引进关于一条闭曲线对于一个向量场的 指数 的概念.

假定 f 在 xy 平面的集合 R 内是连续的并且有连续的一次微商，那么 f 在 R 的每一点处决定了两个量

$$u = f_x(x, y), \quad v = f_y(x, y). \tag{74}$$

这些量可以看作是一个确定的向量——f 的 梯度 的分量. R 内各点处的梯度形成一个 向量场. R 内的临界点就是梯度等于零的那些点. 在所有其他点处，梯度向量都有一个唯一确定的方向，例如用它的 方向余弦

$$\xi = \frac{u}{\sqrt{u^2 + v^2}} \quad 与 \quad \eta = \frac{v}{\sqrt{u^2 + v^2}}$$

所确定的方向. 显然，ξ 与 η 在每一个非临界点处是 (x, y) 的连续函数. 我们可以令

$$\xi = \cos\theta, \qquad \eta = \sin\theta,$$

不过其中的角 θ——向量 (u, v) 的 倾角——除去 2π 的一个整倍数以外才是确定的. 一般地说，对于 θ 要选择一个确定的值使其随 (x, y) 连续变化这是不可能的. 另一方面，微分

$$\begin{aligned} d\theta &= d\arctan\frac{v}{u} = \frac{udv - vdu}{u^2 + v^2} \\ &= \frac{(uv_x - vu_x)dx + (uv_y - vu_y)dy}{u^2 + v^2} \end{aligned} \tag{75}$$

对于 R 的每一个非临界点都是明确地确定了的.

现在设 C 是 R 内一条定向的闭曲线, 而且不经过 f 的任何一个临界点. 我们定义数

$$I_c = \frac{1}{2\pi}\int_C d\theta = \frac{1}{2\pi}\int_C \frac{udv - vdu}{u^2+v^2} \tag{76}$$

为曲线 C 对于向量场的 **庞加莱指数**. 如果 C 由参数形式

$$x = \phi(t), y = \psi(t) \quad (a \le t \le b)$$

给出, 其中 ϕ 与 ψ 在定义区间的两端有同样的值, 而且曲线 C 的指向是对应于 t 增加的方向, 那么 C 的指数由积分

$$I_c = \frac{1}{2\pi}\int_a^b \left(\frac{u}{u^2+v^2}\frac{dv}{dt} - \frac{v}{u^2+v^2}\frac{du}{dt}\right)dt$$

给出. 因为当沿着曲线 C 走几圈后回到同一点 (x,y) 时, 对应于 $t=a$ 与 $t=b$ 的 θ 值只差一个 2π 的整倍数, 所以 I_C 总是一个整数. 这个整数算出当我们沿曲线 C 指定方向行进时, 向量 (u,v) 反时针方向旋转的圈数 [1]. 当然, 如果 C 改变定向, 那么 I_C 也就变号. 作为一个例证, 考虑函数

$$f(x,y) = x^2 + y^2.$$

这里梯度

$$(u,v) = (2x, 2y)$$

在每一点 (x,y) 处的方向就是从原点到该点的向径的方向. 假定我们用的是右手坐标系. 对于一条不经过原点的闭曲线 C, 指数

$$I_c = \frac{1}{2\pi}\int_C \frac{xdy - ydx}{x^2+y^2}$$

就给出从原点出发的向径沿曲线 C 反时针旋转的圈数. 这实际上就是第一卷 (第 489 页) 中所证明的曲线 C 绕原点的次数的公式.

1) 为了定义 "指数", 向量场不必一定是梯度场.

一般说来，在 u 与 v 不同时为零的那些点处，方程 (75) 的微分 $d\theta$ 满足可积条件

$$\left(\frac{uv_x - vu_x}{u^2+v^2}\right)_y = \left(\frac{uv_y - vu_y}{u^2+v^2}\right)_x,$$

这个条件可以直接验证，而且它所反映的不过是这样一个关系

$$\left[\left(\arctan\frac{v}{u}\right)_x\right]_y = \left[\left(\arctan\frac{v}{u}\right)_y\right]_x,$$

这个关系不论函数 $\arctan(v/u)$ 的多值性也总是成立的. 这样就**可由关于线积分的基本定理**(见第 113 页与第 105 页)**推知: 如果 C 是 R 的一个不包含 f 的临界点的单连通子集的边界, 那么** $I_C = 0$.

更一般地说，考虑具有数条闭的边界曲线 $C_1, C_2, \cdots, C_n$ 的多连通集合 R. 设 (x, y) 坐标系为通常的右手系. 假定每一条曲线 C_i 的方向确定如下：当我们在 C_i 上按其指向行进时，R 在我们的左边. 假设我们可以用适当的补助弧线连结各个 C_i (见图 3.31) 把 R

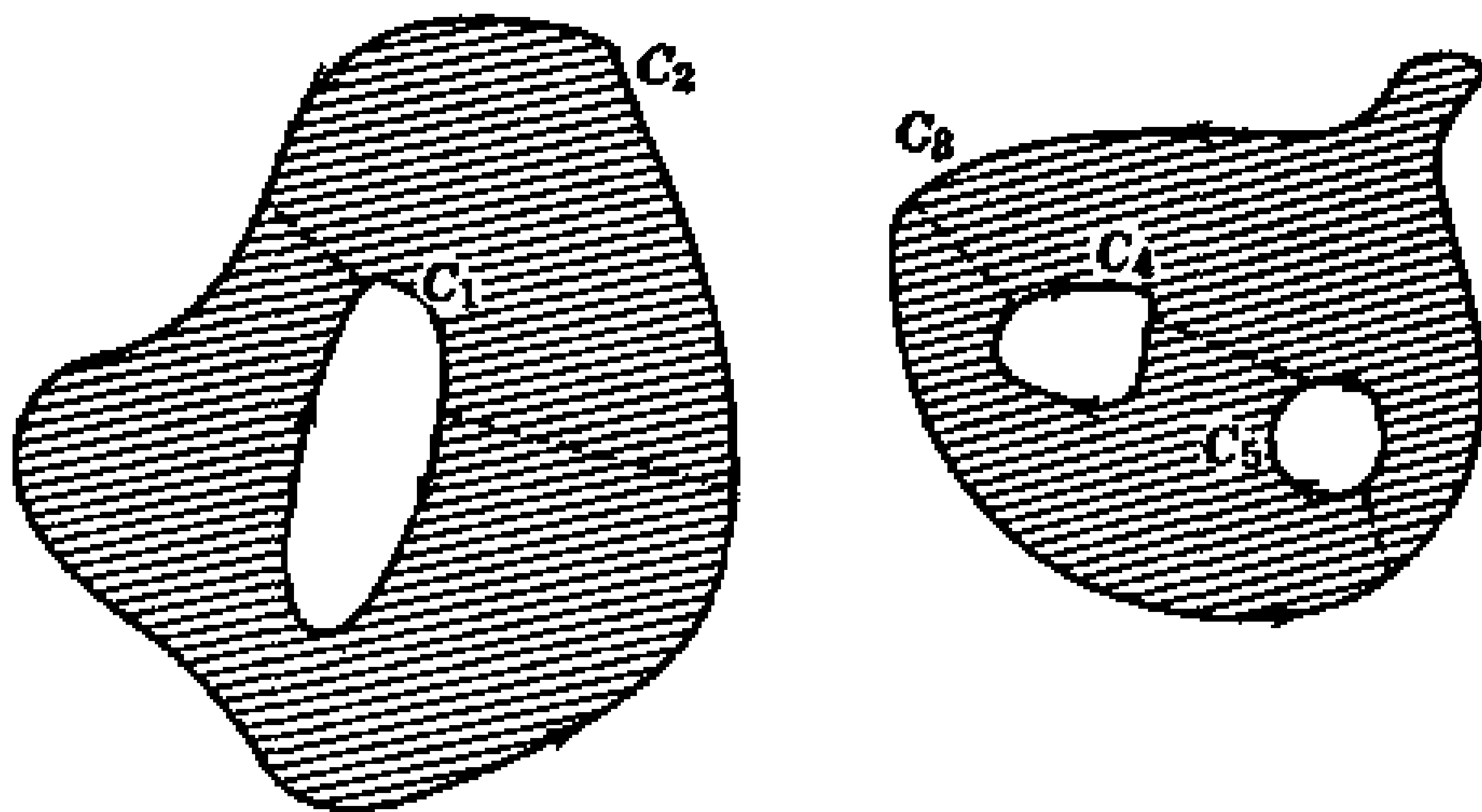


图 3.31　具有正的定向边界曲线 C_i 的多连通区域被分成单连通集合

分割成为一些单连通集合 R_k. 设 f 在 R 内没有临界点，那么

$$\int d\theta = 0.$$

这里积分路线是任何一个 R_k 的边界沿反时针方向. 把对所有 R_k 的边界的积分组成和数，我们看到补助弧线上的积分消掉了，因而

我们得到

$$0 = \sum_i \int_{C_i} d\theta$$

而这意味着

$$\sum_{i=1}^{n} I_{C_i} = 0, \tag{77}$$

其中 C_i 是闭曲线组成一个没有 f 临界点的集合 R 的边界，且定向使 R 在其左边.

作为一个结论我们得到这一定理：如果 R 的边界曲线 (定向如上所说) 的指数和不等于零，那么 R 内至少有一个临界点.

关于 R 内临界点的个数的更为确切的结果是可以得到的，如果我们假定 f 在 R 内有连续的二阶微商，f 只有有限个数的临界点 $(x_1, y_1), \cdots, (x_N, y_N)$，并且在每个临界点处判别式

$$D = f_{xx}f_{yy} - f_{xy}^2$$

不等于零. 这时所有的临界点不是极大或极小 (对应于 $D > 0$) 就是鞍点 (对应于 $D < 0$)(见第 383 页). 仍假定 R 是被定向的简单闭曲线 $C_1, C_2, \cdots, C_n$ 所围成的，这些曲线不通过 f 的任何临界点. 我们可以挖掉每个临界点 (x_k, y_k) 用曲线 γ_k 围成的一个小邻域. 这样剩下的由曲线 $C_1, \cdots, C_n, \gamma_1, \cdots, \gamma_N$ 所围成的一个集合就没有 f 的临界点. 给每一条 γ_k 以反时针的方向，按 (77) 我们就有

$$\sum_{i=1}^{n} I_{C_i} - \sum_{k=1}^{N} I_{\gamma_k} = 0. \tag{78}$$

现在每一条曲线 γ_k 所围成的集合内只有一个临界点，所以 γ_k 的指数只依赖于该临界点的类型，这就是我们将要证明的.

设 γ_k 是中心在临界点 (x_k, y_k) 而半径为 r 的一个小圆

$$x = x_k + r\cos t, \quad y = y_k + r\sin t.$$

根据泰勒定理，在 γ_k 上我们有

$$u=f_x(x,y)=(x-x_k)f_{xx}(x_k,y_k)+(y-y_k)f_{xy}(x_k,y_k)+\cdots$$

$$=r(a\cos t+b\sin t)+O(r^2), \tag{79a}$$

$$v=f_y(x,y)=(x-x_k)f_{yx}(x_k,y_k)+(y-y_k)f_{yy}(x_k,y_k)+\cdots$$

$$=r(b\cos t+c\sin t)+O(r^2), \tag{79b}$$

其中我们令

$$a=f_{xx}(x_k,y_k),\ b=f_{xy}(x_k,y_k),\ c=f_{yy}(x_k,y_k).$$

为了求出 t 由 0 变到 2π 时向量 (u,v) 作逆时针方向旋转的次数，我们注意到平面上以 (u,v) 为坐标的点 (即这个点的位置向量的分量是 u,v) 近似地描出具有参数表示

$$u=r(a\cos t+b\sin t),\ v=r(b\cos t+c\sin t) \tag{80}$$

的椭圆 E. 这个椭圆以原点为中心并且有非参数方程

$$(cu-bv)^2+(av-bu)^2=r^2(ac-b^2)^2.$$

很明显，当 t 从 0 增加到 2π 时，点 (u,v) 恰好把 (80) 式中的椭圆 E 描一圈，因此 γ_k 的指数不是 $+1$ 就是 -1, 这分别对应当 t 增加时 E 是反时针或顺时针方向. 现在很清楚，线性映射

$$u=r(au+bv),\quad v=r(bu+cv)$$

把在 uv 平面的圆 (其中 t 增加的方向对应于圆的反时针方向)

$$u=\cos t,\quad v=\sin t$$

映射成 E. 由于曲线的方向保持或是反转对应于映射的雅可比行列

式 $r^2(ac-b^2)$ 的符号，我们看到 [1]

$$I_{r_k}=\operatorname{sgn}(ac-b^2)=\operatorname{sgn}[f_{xx}(x_k,y_k)f_{yy}(x_k,y_k)-f_{xy}^2(x_k,y_k)]$$

$$=\operatorname{sgn}D(x_k,y_k)$$

由 (78) 式就得到

$$\sum_{i=1}^{n}I_{C_i}=\sum_{k=1}^{N}\operatorname{sgn}D(x_k,y_k).$$

正如我们看到的，当临界点 (x_k,y_k) 是极大点或极小点时 $\operatorname{sgn}D(x_k,y_k)=+1$; 而当它是鞍点时 $\operatorname{sgn}D(x_k,y_k)=-1$. 设 M_0,M_1,M_2 分别表示 R 内极小点、鞍点与极大点的个数，我们的结果就成了庞加莱恒等式 [2]

$$\sum_{i=1}^{n}I_{C_i}=M_0-M_1+M_2. \tag{81}$$

总之，f 在 R 内的极大点与极小点的总数超过鞍点总数的差额等于 R 的边界曲线对于 f 的梯度场的指数的和数，其中每条边界曲线的定向使得 R 在其左侧.

当 f 在 R 的每一条边界曲线 C_i 上为常数时，这个结果特别简单. 这时 f 的梯度向量垂直于 C_i(见 253 页) 并且具有 C_i 的外法线或内法线的方向. 如果 C_i 上没有 f 的临界点并且 C_i 是一条光滑的简单闭曲线，那么梯度的方向就连续变化并且在 C_i 的任何一点处都不可能从外法向跳到内法向或从内法向跳到外法向. 于是很

1) 这个结果也可以用分析方法得到，只需注意，按公式 (79a,b) 有

$$\lim_{r\to 0}I_{r_k}=\lim_{r\to 0}\frac{1}{2\pi}\int_{r_k}\frac{udv-vdu}{u^2+v^2}$$

$$=\frac{1}{2\pi}\int_0^{2\pi}\frac{(ac-b^2)dt}{(a\cos t+b\sin t)^2+(b\cos t+c\sin t)^2}.$$

这个积分可以算出 (见第一卷第 329 页)，其值为 $2\pi\operatorname{sgn}(ac-b^2)$.

2) 多于两个自变量的函数的相应的公式就是莫尔斯公式.

明显，梯度向量沿着 C_i 完整地转一圈，与梯度向量具有固定夹角的 C_i 的切线向量也沿同样方向转一圈. 因此当 C_i 具有反时针方向时，$I_{C_i}=+1$; 当 C_i 具有顺时针方向时，$I_{C_i}=-1$. 容易看出，根据我们关于 R 的边界曲线的定向的约定，一条边界曲线 C_i, 当它是组成 R 的一个不连通部分的外边界时，它具有反时针方向; 当它围成 R 内的一个"洞 (hole)"时，它具有顺时针方向 (见图 3.31). 由此推知，对于在边界曲线上为常数的 f, 有恒等式

$$M_0 - M_1 + M_2 = N_0 - N_1, \tag{82}$$

其中 N_0 是 R 的连通部分的块数，N_1 是 R 内全部洞的个数 (R 的"连通数").

作为例子，取 R 是圆盘的情形. 这里 $N_0=1, N_1=0$, 因此对于在边界上为常数的 f 便有

$$M_0 - M_1 + M_2 = 1.$$

我们在这里求得 *R 的内部临界点的总数是*

$$M_0 + M_1 + M_2 = 1 + 2M_1,$$

这是一个奇数. 此外, 如果 f 的极值点的个数 M_0+M_2 超过1, 那么在 R 内至少有一个鞍点.

对于一个圆环，我们有

$$N_0 = 1, \qquad N_1 = 1,$$

因此，对于在每一条边界曲线上为常数的 f 便有

$$M_0 - M_1 + M_2 = 0.$$

试看 f 在两条边界曲线上有同一常数值的情形. 这时 f 或者处处是常数或者在 R 的内部达到它的最大或最小. 如果我们假定 f 只

有一个使得 $f_{xx}f_{yy}-f_{xy}^2\neq 0$ 的临界点，f 为常数的情况就被排除了．由此得到 $M_0+M_1>0$，于是 $M_1>0$．因此，**一个在圆环上的函数，如果它在边界上处处等于零，那么它在圆环内部至少有一个使 $f_{xx}f_{yy}-f_{xy}^2\leq 0$ 的临界点．**

练　习　A.2

1. 试举一个原点是奇点的连续函数 f 的例子，使其奇点的指数为：

(a) -1;

(b) -2;

(c) $-n$，其中 n 是一个自然数．

2. 试举一个原点是奇点的函数 f 的例子，不要求它是连续的，但使其奇点的指数为：

(a) 2;

(b) n，其中 n 是一个自然数．

3. 设 R 是 xy 平面内由一条具有连续转动切线的闭的凸曲线围成的凸区域；又设

$$\xi=f(x,y),\quad \eta=g(x,y)$$

是一个 R 到其自身的连续可微映射．试证：R 内至少有一个"不动点"，即 R 内存在一点 (x,y) 使得

$$x=f(x,y),\quad y=g(x,y).$$

n 维空间中类似的不动点定理是属于布劳威尔 (Brouwer) 的．[提示：考虑分量为 $u=f(x,y)-x$，$v=g(x,y)-y$ 的向量场．]

A.3　平面曲线的奇点

在第 256 页上我们看到，通常一条曲线 $f(x,y)=0$ 在一点

$(x,y)=(x_0,y_0)$ 是一个奇点，当在这点有这样三个方程

$$f(x_0,y_0)=0,\quad f_x(x_0,y_0)=0,\quad f_y(x_0,y_0)=0$$

成立．为了系统地研究这些奇点，我们假定函数 $f(x,y)$ 在 (x_0,y_0) 的邻域内有直到二阶的连续微商，并且二次微商在该点不全为零．按泰勒级数展开到二阶项，我们得到曲线的方程形如

$$\begin{aligned}2f(x,y)=&(x-x_0)^2f_{xx}(x_0,y_0)+2(x-x_0)(y-y_0)f_{xy}(x_0,y_0)\\&+(y-y_0)^2f_{yy}(x_0,y_0)+\varepsilon\rho^2=0,\end{aligned}$$

其中 $\rho^2=(x-x_0)^2+(y-y_0)^2$，且当 ρ 趋于零时 ε 趋于零．

应用参量 t，我们可以把通过点 (x_0,y_0) 的一般直线方程写成这样的形式：

$$x-x_0=at,\qquad y-y_0=bt,$$

其中 a 与 b 是两个任意常数，我们可以假定是按 $a^2+b^2=1$ 选取的．为了确定这条直线与曲线 $f(x,y)=0$ 的交点，我们把这两个表达式代入上面对于 $f(x,y)$ 的展开式．这样，对于交点，我们就得到方程

$$a^2t^2f_{xx}+2abt^2f_{xy}+b^2t^2f_{yy}+\varepsilon t^2=0.$$

第一个解是 $t=0$，即点 (x_0,y_0) 本身，这是显然的．值得注意的是，方程的左边是可以被 t^2 除尽，因此 $t=0$ 是方程的一个**二重根**．根据这个理由，奇点有时候也称为曲线的**二重点**．如果我们约去因子 t^2，便得到方程

$$a^2f_{xx}+2abf_{xy}+b^2f_{yy}+\varepsilon=0.$$

我们现在讨论前面说的那条曲线是否可能与这曲线有别的交点，该点当这直线趋于某个特定位置时趋于 (x_0,y_0)．一条割线的这样一个极限位置，我们当然叫做一条切线．要讨论这个问题，我们注意到，当一个点趋于 (x_0,y_0) 时，量 t 趋于零，因之 ε 也趋于

零. 如果上述方程仍然满足, 那么表达式 $a^2f_{xx}+2abf_{xy}+b^2f_{yy}$ 必定也趋于零, 即对于直线的极限位置, 我们必定有

$$a^2f_{xx}+2abf_{xy}+b^2f_{yy}=0.$$

这个方程给了我们一个二次条件来确定比值 a/b, 而比值 a/b 确定切线的斜率.

如果方程的判别式是负的, 就是说, 如果

$$f_{xx}f_{yy}-f_{xy}^2<0,$$

我们就得到两条不同的切线. 这时曲线有一个二重点或结点, 如双纽线 $(x^2+y^2)^2-(x^2-y^2)=0$ 在原点或环索线

$$(x^2+y^2)(x-2a)+a^2x=0$$

在 $(x_0,y_0)=(a,0)$ 这点所显示的一样.

如果判别式为零, 即如果

$$f_{xx}f_{yy}-f_{xy}^2=0,$$

我们就得到两条重合的切线; 这时可能是曲线的两个分支彼此相切或曲线有一个尖点[1].

最后, 如果

$$f_{xx}f_{yy}-f_{xy}^2>0,$$

就根本没有(实的)切线. 这种情形出现在, 例如一条代数曲线的称为孤立点的情形, 在这样的点上, 曲线的方程是满足的, 但是在其邻域内再没有曲线的其他的点.

曲线 $(x^2-a^2)^2+(y^2-b^2)^2=a^4+b^4$ 就是一个例子. 数值 $x=0$, $y=0$ 满足方程, 但对于区域

$$|x|<\sqrt{2}a,\qquad |y|<\sqrt{2}b$$

1) 在这种情形, 曲线不一定有一个奇点; 例如 $f(x,y)=(x-y)^2$ 在原点处.

内所有别的值，左边是小于右边的.

我们略去了所有二阶微商都等于零的情形，这种情形过于复杂，所以我们就不讨论了. 通过这样的点可能有曲线的若干个分支，或者可能是其他类型的奇点.

最后，我们简要地提一下这些问题跟最大最小理论之间的联系. 因为一阶微商等于零，所以曲面 $z=f(x,y)$ 在稳定点 (x_0,y_0) 的切平面方程就是

$$z-f(x_0,y_0)=0.$$

于是方程

$$f(x,y)-f(x_0,y_0)=0$$

给出切平面与曲面的交线在 xy 平面上的投影曲线，因而我们看到点 (x_0,y_0) 是这条曲线的一个奇点. 如果这是一个孤立点，那么在其某一邻域内切平面与曲面没有别的公共点，**并且**函数 $f(x,y)$ 在 (x_0,y_0) 这点有一个极大或极小 (参看第 385 页). 然而，如果这奇点是一个多重点，切平面就与曲面相交成具有两个分支的曲线，因而 (x_0,y_0) 就是一个鞍点. 这些讨论恰好把我们引导到我们早已在 A.1 节中就求得了的充分条件.

练　习　A.3

1. 试求下列各曲线的奇点，并讨论其性质:

(a) $(x^2+y^2)^2-2c^2(x^2-y^2)=0\ (c\neq 0)$,

(b) $x^2+y^2-2x^3-2y^3+2x^2y^2=0$,

(c) $x^4+y^4-2(x-y)^2=0$,

(d) $x^5-x^4+2x^2y-y^2=0$.

A.4　曲面的奇点

用类似的方法我们可以讨论曲面 $f(x,y,z)=0$ 的奇点，也就

是适合条件

$$f=0,\ f_x=f_y=f_z=0$$

的点. 不失一般性, 我们可以把这个点取为原点. 如果我们把该点的值记为

$$f_{xx}=\alpha,\ f_{yy}=\beta,\ f_{zz}=\gamma,\ f_{xy}=\lambda, f_{yz}=\mu,\ f_{xz}=\nu,$$

我们就得到方程

$$\alpha x^2+\beta y^2+\gamma z^2+2\lambda xy+2\mu yz+2\nu xz=0$$

这里点 (x,y,z) 是在曲面在 O 点处的切平面上.

这个方程表示一个在奇点与曲面相切的二次锥面 (它代替了曲面在正常点上的切平面). 如果我们假定量 $\alpha,\beta,\cdots,\nu$ 不全为零, 并且假定上述方程有不同于 $x=y=z=0$ 的实数解.

练 习 A.4

1. 应用 A.1 节中习题 6 的结果, 考察曲面在奇点的一个邻域内的性态.

A.5 流体运动的欧拉表示法与拉格朗日表示法之间的联系

设 (a,b,c) 是流体 (液体或气体) 内的一个质点在时刻 $t=0$ 时的坐标. 于是这质点的运动可以用三个函数

$$x=x(a,b,c,t),$$

$$y=y(a,b,c,t),$$

$$z=z(a,b,c,t),$$

或者一个位置向量 $\mathbf{X}=\mathbf{X}(a,b,c,t)$ 来表示. 速度与加速度由对于 t 的微商所给出: 速度向量是 $\dot{\mathbf{X}}$, 其分量为 $\dot{x},\dot{y},\dot{z}$; 加速度向量是 $\ddot{\mathbf{X}}$, 其分量为 $\ddot{x},\ddot{y},\ddot{z}$. 所有这些量都是作为初始位置 (a,b,c) 及参量 t 的函数出现的. 对于 t 的每一个值, 我们有一个变换, 它把属于流体内不同的点的坐标 (a,b,c) 变换成时刻 t 的坐标 (x,y,z). 这就是所谓 **运动的拉格朗日表示法**. 由欧拉引进的另一表示法是基于把三个函数

$$u(x,y,z,t),\ v(x,y,z,t),w(x,y,z,t)$$

理解为 (x,y,z) 点在时刻 t 的速度 $\dot{\mathbf{X}}$ 的三个分量 $\dot{x},\dot{y},\dot{z}$.

为了从第一种表示法过渡到第二种表示法, 我们必须用第一种表示法, 把 a,b,c 解成 x,y,z 和 t 的函数, 并把这些表达式代入表达式 $\dot{x}(a,b,c,t)$, $\dot{y}(a,b,c,t)$, $\dot{z}(a,b,c,t)$:

$$u(x,y,z,t)=\dot{x}(a(x,y,z,t),b(x,y,z,t),c(x,y,z,t),t),\cdots.$$

我们然后从

$$\dot{x}(a,b,c,t)=u(x(a,b,c,t),y(a,b,c,t),z(a,b,c,t),t),\cdots$$

关于固定的 a,b,c 而对于 t 求微商, 得到加速度的分量

$$\ddot{x}=u_x\dot{x}+u_y\dot{y}+u_z\dot{z}+u_t,\cdots,$$

或

$$\ddot{x}=u_xu+u_yv+u_zw+u_t,$$

$$\ddot{y}=v_xu+v_yv+v_zw+v_t,$$

$$\ddot{z}=w_xu+w_yv+w_zw+w_t.$$

在流体力学中, 下述连接欧拉表示法与拉格朗日表示法的方程是基本的

$$\operatorname{div}\dot{X}=u_x+v_y+w_z=\frac{\dot{D}}{D},$$

其中

$$D(x,y,z,t)=\frac{d(x,y,z)}{d(a,b,c)}$$

是刻画变换的特性的雅可比行列式.

读者可以用隐函数微分法 (见第 275 页) 的各种法则以及对于二维空间的相应的定理来完成这件事的证明.

练 习 A.5

1. 关系式 $u_t=v_t=w_t=0$ 的物理意义是什么?
2. 对关系式

$$\ddot{x}=u_x u+u_y v+u_z w+u_t,$$

$$\ddot{y}=v_x u+v_y v+v_z w+v_t,$$

$$\ddot{z}=w_x u+w_y v+w_z w+w_t$$

作出物理解释, 并用向量记号表示这些关系.

A.6 闭曲线的切线表示法与周长不等式

带有参量 α 的一个直线族可以由

$$x\cos\alpha+y\sin\alpha-p(\alpha)=0 \tag{83}$$

给出, 其中 $p(\alpha)$ 是一个函数, 它二次连续可微并且以 2π 为周期 (这里 p 表示原点到直线族中具有法方向 α 的那条直线的距离). 这些直线的包络 C 是一条闭曲线, 它满足 (83) 以及另一方程

$$-x\sin\alpha+y\cos\alpha-p'(\alpha)=0.$$

因此

$$x=p\cos\alpha-p'\sin\alpha,$$

$$y = p\sin\alpha + p'\cos\alpha \tag{84}$$

是 C 的参量表示式 (α 为参量). 公式 (83) 给出了 C 的切线表示式, 称为 C 的 **切线方程** [1], 而 $p(\alpha)$ 称为 C 的 **支撑函数**.

因为

$$x' = -(p+p'')\sin\alpha, \quad y' = (p+p'')\cos\alpha,$$

所以我们立刻得到下面关于 C 的长度 L 与面积 A 的表示式:

$$L = \int_0^{2\pi} \sqrt{x'^2+y'^2}d\alpha = \int_0^{2\pi}(p+p'')d\alpha = \int_0^{2\pi} pd\alpha,$$

$$\begin{aligned} A &= \frac{1}{2}\int_0^{2\pi}(xy'-yx')d\alpha = \frac{1}{2}\int_0^{2\pi}(p+p'')pd\alpha \\ &= \frac{1}{2}\int_0^{2\pi}(p^2-p'^2)d\alpha, \end{aligned}$$

因为 $p'(\alpha)$ 也是 2π 为周期的函数 [2].

由此我们可导出这 **等周不等式**

$$L^2 \geq 4\pi A,$$

其中等号只对于圆周成立. 这也可以这样陈述: 在给定长度的全部闭曲线中, 圆周包围着最大的面积.

为了证明, 我们利用 $p(\alpha)$ 的傅里叶展开式

$$p(\alpha) = \frac{a_0}{2} + \sum_{\nu=1}^{\infty}(a_\nu\cos\nu\alpha + b_\nu\sin\nu\alpha);$$

从而

$$p'(\alpha) = \sum_{\nu=1}^{\infty}\nu(b_\nu\cos\nu\alpha - a_\nu\sin\nu\alpha),$$

---

1) 表示法 (84) 适用于任何一条闭的凸曲线 C, 只要 C 的曲率是有限的、正的、沿着 C 连续变化的.

2) 因为 $p(\alpha)+c$ 是与 C 平行且相距为 c 的曲线的支撑函数, 因此一条平行曲线的长度与面积的公式是容易从这些式子得出的 (参考第一卷第 488 页习题 7, 它的解见 A. Blank 的书 Problems in Calculus and Analysis).

所以 我们有

$$L = \pi a_0,$$

$$A = \frac{\pi}{2}\left(\frac{a_0^2}{2} - \sum_{\nu=2}^{\infty}(\nu^2 - 1)(a_\nu^2 + b_\nu^2)\right).$$

因此

$$A \leq \frac{\pi a_0^2}{4} = \frac{L^2}{4\pi};$$

特别是, $A = \dfrac{L^2}{4\pi}$ 仅当 $a_\nu = b_\nu = 0\ (\nu \geq 2)$ 成立; 就是说仅当

$$p(\alpha) = \frac{a_0}{2} + a_1 \cos\alpha + b_1 \sin\alpha.$$

而这个方程定义一个圆, 可从 (84) 立即得到证明.

练 习 A.6

1. 就下列各族直线求出包络, 及其长度以及所包围着的面积:

(a) $(x+2)\cos\alpha + y\sin\alpha + 2 = 0$,

(b) $x\cos\alpha + y\sin\alpha + \dfrac{1}{2}\sin 2\alpha = 0$.

2. 比较面积公式与长度公式. 可能存在任意长的曲线而围有任意小的面积吗?

3. 每一条闭曲线都能表示成 (83) 那样的直线族的包络吗?

解　答

第一章

练习 1.1 (p.11)

1. (a) 用极坐标形式写出 $z = r(\cos\theta + i\sin\theta)$, $0 < \theta < 2\pi$. 则根据棣莫弗定理 有

$$z^n = r^n(\cos n\theta + i\sin n\theta).$$

对于 $r < 1$, 我们有 $\lim\limits_{n\to\infty} r^n = 0$; 因此, $\lim\limits_{n\to\infty} z^n = 0$. 对于 $r > 1$, 我们有 $\lim\limits_{n\to\infty} r^n = \infty$; 因此, z^n 到原点, 从而到任何给定点的距离都能变得任意大, 于是这个序列发散. 对于 $r = 1$, 有两种情况: $z = 1(\theta = 0)$, 对此有 $\lim\limits_{n\to\infty} z^n = 1$, 及 $z = \cos\theta + i\sin\theta$. 在后一种情况, 对于这序列的相继两点之间距离我们有

$$|z^{n+1} - z^n| = |z^n| \cdot |z - 1| = |z - 1| = \sqrt{2 - 2\cos\theta},$$

这是一个固定的正值; 于是根据柯西准则, 这个序列一定要发散.

(b) 复数 z 的 n 次方根是用极坐标形式这样来给定的:

$$z^{1/n} = r^{1/n}\left(\cos\frac{\theta}{n} + i\sin\frac{\theta}{n}\right).$$

如果 $z = 0$, 我们有 $\lim\limits_{n\to 0} z^{1/n} = 0$. 否则, 令 $z^{1/n} = x_n + iy_n$, 便有

$$\begin{aligned}\lim_{n\to\infty} z^{1/n} &= \lim_{n\to\infty} x_n + i\lim_{n\to\infty} y_n \\ &= \lim_{n\to\infty} r^{1/n}\cos\frac{\theta}{n} + i\lim_{n\to\infty} r^{1/n}\sin\frac{\theta}{n} = 1.\end{aligned}$$

2. 将第一卷中所有的极限定理分别应用到 P_m 的分量上.

3. 对于满足 $a^2+b^2<1$ 的一个点 (a,b), 令

$$\alpha=\sqrt{a^2+b^2}.$$

则 (a,b) 的邻域 $(x-a)^2+(y-b)^2<(1-\alpha)^2$ 含于单位圆内.

对于满足 $a^2+b^2=1$ 的一个点 (a,b), 每个邻域都包含着不在单位圆内的点.

4. 设 (a,b) 是 S 的任一点, 令 $\gamma=b-a^2>0$. 考察 (a,b) 的一个 ε 邻域,

$$(x-a)^2+(y-b)^2<\varepsilon^2.$$

对这邻域内的一切点, 我们有 $|x-a|<\varepsilon$, $|y-b|<\varepsilon$. 利用

$$a^2=x^2-2(x-a)\cdot a-(x-a)^2,$$

我们得到

$$\begin{aligned}y&>b-\varepsilon=a^2+\gamma-\varepsilon\\&=x^2-2(x-a)a-(x-a)^2+\gamma-\varepsilon\\&>x^2+\gamma-2\varepsilon|a|-\varepsilon^2-\varepsilon>x^2,\end{aligned}$$

只要 ϵ 取为 1 和 $\gamma/(2|a|+2)$ 中较小者. 这样这个 ϵ 邻域就在 S 内.

5. 线段 (如果端点不作为这线段的点, 就把它们同这线段一起作答案).

问题 1.1 (p.12)

1. 按定义, 每一个边界点 P 的邻域都包含 S 的点. 在 S 内取 P_1 使得 $\overline{P_1P}<1/2$. 因为 P 不在 S 内, 就有 $P_1\neq P$, 因此, $\overline{P_1P}>0$. 现在用数学归纳法继续进行: 给定 P_n 在 S 内, 就取 P_{n+1} 使得

$$\overline{P_{n+1}P}<\frac{1}{2}\overline{P_nP}.$$

显然，P_n 是不相重的并且 $\overline{P_nP} < 1/2^n$.

2. 设给定的集合为 S, S 的闭包为 S_c; 以及 S_c 的闭包为 S_{cc}. 每个属于 S_{cc} 的点不在 S_c 内就在 S_c 的边界上. 如果 P 在 S_c 的边界上，那么 P 的每一个邻域至少包含 S_c 内的一个点 Q 和不在 S_c 内的一个点 R. 因为 R 不在 S_c 内，所以就不在 S 内. 又因为每一个邻域都是开集，所以 P 的邻域包含着必有 S 的点含在其中的 Q 的一个邻域. 于是 P 在 S_c 内.

3. 设 X 是 S 的在 $\overline{PQ}$ 上的任一点，因为 $\overline{PX} \le \overline{PQ}$, $\overline{PX}$ 值的集合是有界的. 记它的最小上界为 $\mathrm{lub}\overline{PX}$. 设 R 是在 $\overline{PQ}$ 上与 P 距离等于 $\mathrm{lub}\,\overline{PX}$ 的点，则 R 的任何邻域包含 $\overline{PQ}$ 上属于 S 的点，又包含不属于 S 的点.

4. G 的一切点都是内点.

练习 1.2 (p.17)

1. (a) $\dfrac{27}{8}$; (c) $\dfrac{1}{(\log\pi)^e}$; (e) 5.

2. 定义域是点 (x,y) 的集合，而值域是 u 值的集合，这里

(a) $y \ge -x$, $u \ge 0$; (c) $y > -x$, $u > 0$;

(e) $y > -\dfrac{x}{5}$, u 实数; (g) $x^2+y^2+z^2 \le a^2$, $0 \le w \le a$;

(h) $y \ne -x$, u 实数; (i) $x^2+2y^2 \le 3$, $0 \le u \le \sqrt{3}$;

(j) $x = y = 0$, $u = 0$; (k) $|y| < |x|$, u 实数;

(l) $(x,y) \ne (0,0)$, $0 \le u \le \dfrac{\pi}{4}$. (m) $y \ne -x$, $-\dfrac{\pi}{2} < u < \dfrac{\pi}{2}$;

(n) $x \ne 0$, $0 < u \le 1$; (o) $\dfrac{1}{e} < x+y < e$, $0 \le u \le \pi$;

(p) $2n\pi - \dfrac{\pi}{2} \le x \le 2n\pi + \dfrac{\pi}{2}$ 并且 $y \ge 0$, 或

$2n\pi + \dfrac{\pi}{2} \le x \le 2n\pi + \dfrac{3\pi}{2}$ 并且 $y \le 0$, $\quad u \ge 0$.

3. 对于 k 个变量是

$$\frac{1}{k!}(n+1)(n+2)\cdots(n+k).$$

(参看第一卷第一章， p.133, 练习第 11 题)

练习 1.3 (p.26)

2. 在 $x=y=0$ 不连续.

3. (a) 设 $x=\rho\cos\theta$, $y=\rho\sin\theta$, 则

$$|f(x,y)|=\rho^3|\cos^3\theta-3\cos\theta\sin^2\theta|<4\rho^3.$$

取 $\delta(\varepsilon)=\sqrt[3]{\varepsilon/4}$, 则 $f(x,y)$ 至少有 ρ^3 的阶.

4. 同一元实变量函数理论一样, 连续函数的和与积以及连续函数的连续函数都是连续的.

(a) 连续.

(b) 不连续性只可能出现在 (0,0). 注意到如果引入

$$x=\rho\cos\theta,\qquad y=\rho\sin\theta,$$

则根据 $|\sin\alpha|<|\alpha|$ 便有

$$\left|\frac{\sin xy}{\sqrt{x^2+y^2}}\right|<\rho;$$

因此, 在点 (0,0) 极限存在并且是 0.

5. 对于 $z\geq 0$, $z+h>0$ 用微分中值定理得到

$$|\sqrt{1+(z+h)}-\sqrt{1+z}|=\frac{|h|}{2\sqrt{1+(z+\theta h)}}\leq\frac{|h|}{2};$$

因此, 在 z 适当选取的各种情况下, 要求 $|h|<2\varepsilon$ 就足够了. 令 $\Delta x=\rho\cos\theta$, $\Delta y=\rho\sin\theta$, 这里 $\rho<\delta(\varepsilon,x,y)$.

(a) 在 $z=x^2+2y^2$ 且 $h=\Delta z$ 的情况下, 注意

$$|\Delta z|=\rho|2x\cos\theta+4y\sin\theta+\rho(\cos^2\theta+2\sin^2\theta)|$$

$$\leq\rho(2|x|+4|y|+3\rho)\leq\rho(2|x|+4|y|+3),$$

这里我们利用了 $\delta<1$. 为了 $|\Delta z|<2\varepsilon$, 如下要求就足够了:

$$\delta<\min\left(\frac{2\varepsilon}{2|x|+4|y|+3},1\right).$$

6. 在直线 $y = \pm x$ 上.

7. 在直线 $x = n + \frac{1}{2}, y = n + \frac{1}{2}$ 上.

8. 对于一切值. (按定义, 一个函数在它的定义域外部总是连续的.)

9. 设 $z = 1/u$, 这里 $u = 1 - x^2 - y^2$. 我们有 $|\Delta z| = |\Delta u|/(u + \theta\Delta u)^2$. 对于 $u > 0$, 选取 $|\Delta u| < u/2$, 则 $u + \theta\Delta u > u/2$ 并且

$$|\Delta z| < \frac{4|\Delta u|}{u^2}.$$

现在, 用 $\Delta x = \rho\cos\theta, \Delta y = \rho\sin\theta, \rho < \delta \leq 1$ 并且 $|x|, |y| < 1$, 我们有

$$|\Delta u| = |\rho(2x\cos\theta + 2y\sin\theta) + \rho^2|$$

$$< \rho(2|x| + 2|y| + 1) < 5\delta.$$

因此, 为了保证 $|\Delta z| < \varepsilon$, 取

$$\delta = \min\left[\frac{\varepsilon}{20}(1 - x^2 - y^2)^2, 1\right].$$

10. 用 $x = \rho\cos\theta, y = \rho\sin\theta$, 我们有

$$P = \rho^2(a\cos^2\theta + 2b\sin\theta\cos\theta + c\sin^2\theta) = \rho^2 f(\theta).$$

对于 θ 的任何值表达式 $f(\theta)$ 必定不等于 0, 从而我们应当有

$$ac - b^2 > 0.$$

11. 全不连续, (a) 在直线 $x = 0$ 上, (c) 在直线 $y = -x$ 上.

12. 对于沿一条直线趋近, 可设 $x = \rho\cos\theta, y = \rho\sin\theta, \theta$ 为一固定值. 为了指出 $f(x,y)$ 的间断性, 使趋近沿抛物线 $x = ay^2$, a 任意, 为了指出 $g(x,y)$ 的间断性, 使趋近沿着圆周

$$\left(x - \frac{1}{2}\right)^2 + y^2 = \frac{1}{4}.$$

13. 对于 (e) 和 (g) 极限存在. 对于 (h) 令 $y = e^{-a/|x|}$, a 是任意正数, 并指出由于

$$f(x,y) = \frac{y^{|x|}\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}+\left|\frac{y}{x}\right|},$$

所以 $\lim\limits_{x\to 0} f(x, e^{-a/|x|}) = e^{-a}$.

14. 对于练习第 14 题 (e),

$$\delta(\varepsilon) = -\frac{1}{\sqrt{2}\log\varepsilon}.$$

对于练习第 14 题 (g),

$$\delta = \min\left(-\frac{\log 2}{\log\varepsilon}, \frac{1}{2}\right).$$

15. 首先令 $x = y = 0$, 然后令 $z = 0$.

16. 成立, 因为 $R(x,y)$ 在原点没有定义并且原点是 R 的定义域的一个边界点.

17. (a) 1,

(b) 0,

(c) 0.

18. 令 $y = mx$, 则 $\lim\limits_{x\to 0} z = 3(1-m)/(1+m)$.

19. 参看练习第 13 题.

20. 沿着不同于 $x = 0$ 的任何直线趋近, 都产生极限值 0; 而沿着曲线 $y = a/\log x$ 趋近, 则产生任意的极限值 a.

21. 映射 ϕ 把它的定义域的交在围绕原点半径充分小的圆内的那部分映射到中心在 O 点, 半径为 C_ρ 的区间, 这里常数 C 可以取得与 ρ 无关.

问题 1.3 (p.28)

1. 设 S 是 f 的定义域, S^* 是 f^* 的定义域. 如果 Q 是 S 的一个内点, 那么存在 Q 的一个邻域整个地落在 S 内, 因而对 f^* 的

连续性等同于对 f 的连续性. 如果 Q 在 S^* 内又是 S 的一个边界点, 那么无论 Q 是否在 S 内, 都存在 Q 的一个 δ 邻域, 在其中

$$|\delta(P) - f^*(Q)| < \varepsilon/2.$$

对 S^* 中任一点 $\hat{Q}$, 如果它在 Q 的 δ 邻域内但不在 S 内, 那么存在 S 中的点 P 使得 $f(P)$ 任意接近 $f^*(\hat{Q})$, 也就是说

$$|f(P) - f^*(\hat{Q})| < \varepsilon/2.$$

于是有 $|f^*(\hat{Q}) - f^*(Q)| < \varepsilon$.

2. 如果 $\lim\limits_{(x,y)\to(\xi,\eta)} f(x,y) = L$ 并且 $\lim\limits_{n\to\infty}(x_n, y_n) = (\xi,\eta)$, 则对任一正数 ε, 存在 δ 使得 $|f(x,y) - L| < \varepsilon$ 对一切落在 (ξ,η) 的 δ 邻域内的点 (x,y) 成立. 进一步, 存在 N 使得 (x_n, y_n) 对一切 $n > N$ 落在 (ξ,η) 的 δ 邻域内. 因此, 对于 $n > N$ 有 $|f(x_n, y_n) - L| < \varepsilon$.

反之, 假定对于每一个来自 f 定义域并以 (ξ,η) 为极限的点列 (x_n, y_n) 都有 $\lim\limits_{n\to\infty} f(x_n, y_n) = L$. 假若 f 在 (ξ,η) 不以 L 为极限, 那么对某一个 $\varepsilon > 0$ 和一切 $\delta > 0$, 都存在一个点 (x,y) 属于 (ξ,η) 的 δ 邻域, 但 $(x,y) \neq (\xi,\eta)$, 使得

$$|f(x,y) - L| \geq \varepsilon.$$

令 $\delta_1 = 1$ 并取 (x_1, y_1) 属于 (ξ,η) 的 δ_1 邻域使得

$$|f(x_1, y_1) - L| \geq \varepsilon.$$

接着按顺序定义

$$\delta_n = \frac{1}{2}\sqrt{(x_{n-1} - \xi)^2 + (y_{n-1} - \eta)^2}$$

并取 (x_n, y_n) 使得 $\sqrt{(x_n - \xi)^2 + (y_n - \eta)^2} < \delta_n$ 但

$$|f(x_n, y_n) - L| \geq \varepsilon.$$

用这个方法构造出来的 (x_n, y_n) 与前提矛盾，于是 f 在 (ξ, η) 有极限 L.

练习 1.4 a (p.33)

1. (a) $\dfrac{\partial z}{\partial x} = nax^{n-1}$, $\dfrac{\partial z}{\partial y} = mby^{m-1}$;

(c) $\dfrac{\partial z}{\partial x} = \dfrac{\partial x^2 - 3y^2}{x^2 y}$, $\dfrac{\partial z}{\partial y} = \dfrac{3y^2 - 2x^2}{xy^2}$;

(e) $\dfrac{\partial z}{\partial x} = 2xy^{3/2}$, $\dfrac{\partial z}{\partial y} = \dfrac{3}{2}x^2 y^{\frac{1}{2}}$;

(g) $\dfrac{\partial z}{\partial x} = y^{3/4}/2x^{1/2}$, $\dfrac{\partial z}{\partial y} = \dfrac{3x^{1/2}}{4y^{1/4}}$;

(j) $\dfrac{\partial z}{\partial x} = -2x\sin(x^2 + y)$, $\dfrac{\partial z}{\partial y} = -\sin(z^2 + y)$;

(l) $\dfrac{\partial z}{\partial x} = -\dfrac{\sin x}{\sin y}$, $\dfrac{\partial z}{\partial y} = -\dfrac{\cos x \cos y}{\sin^2 y}$;

(n) $\dfrac{\partial z}{\partial x} = \dfrac{2x^2 + y^2}{\sqrt{x^2 + y^2}}$, $\dfrac{\partial z}{\partial y} = \dfrac{xy}{\sqrt{x^2 + y^2}}$.

2. (a) $\dfrac{\partial f}{\partial x} = \dfrac{2x}{3(x^2 + y^2)^{2/3}}$, $\dfrac{\partial f}{\partial y} = \dfrac{2y}{3(x^2 + y^2)^{2/3}}$;

(c) $\dfrac{\partial f}{\partial x} = e^{x-y}$, $\dfrac{\partial f}{\partial y} = -e^{x-y}$;

(e) $\dfrac{\partial f}{\partial x} = yz\cos xz$, $\dfrac{\partial f}{\partial y} = \sin xz$, $\dfrac{\partial f}{\partial z} = xy\cos xz$;

3. (a) $\dfrac{\partial f}{\partial x} = y$, $\dfrac{\partial f}{\partial y} = x$, $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{\partial^2 f}{\partial y^2} = 0$, $\dfrac{\partial^2 f}{\partial x \partial y} = 1$;

(c) 用 $f(x, y) = \dfrac{x + y}{1 - xy}$,

$$\frac{\partial f}{\partial x} = \frac{1 + y^2}{(1 - xy)^2}, \quad \frac{\partial f}{\partial y} = \frac{1 + x^2}{(1 - xy)^2},$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{2(y + y^3)}{(1 - xy)^3}, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{2(x + y)}{(1 - xy)^3},$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{2(x + x^3)}{(1 - xy)^3};$$

(e) $\dfrac{\partial f}{\partial x}=yx^{y-1}e^{(x^y)},\ \dfrac{\partial f}{\partial y}=x^ye^{(x^y)}\log x,$

$$\frac{\partial^2 f}{\partial x^2}=yx^{y-2}e^{(x^y)}(y-1+yx^y),$$

$$\frac{\partial^2 f}{\partial x\partial y}=x^{y-1}e^{(x^y)}(1+y\log x+yx^y\log x),$$

$$\frac{\partial^2 f}{\partial y^2}=x^y(\log x)^2e^{(x^y)}(1+x^y).$$

4. $f_x=0,\ f_y=0,\ f_z=-3.$

5. 1.

8. $(2/r)$.

9. $a=-3$.

问题 1.4 a (p.34)

1. $\binom{n+k}{k}$. (参看练习 1.2 第 3 题)

2. 考察形如 $f(x,y)=\alpha(x)\beta(y)$ 的函数, 其中 α 是可微的而 β 不是.

3. 对于一切 x 和 y, 关于 x 和 y 求微商得到

$$\phi'(x^2+y^2)=\frac{\psi'(x)}{2x}\psi(y)=\frac{\psi'(y)}{2y}\psi(x);$$

从而 $\psi'(x)/2x\psi(x)$ 是常数, $f(x,y)=ce^{a(x^2+y^2)}$.

练习 1.4 c (p.39)

2. (a) 注意观察一阶偏微商

$$\frac{\partial f}{\partial x}=\begin{cases}\dfrac{2x}{(x^2+y^2)^2}\exp[-1/(x^2+y^2)], & x,y\neq 0,\\ 0, & x=y=0,\end{cases}$$

$$\frac{\partial f}{\partial y}=\begin{cases}\dfrac{2y}{(x^2+y^2)^2}\exp[-1/(x^2+y^2)], & x,y\neq 0,\\ 0, & x=y=0,\end{cases}$$

二者都是有界的.

(b) 只有原点是有问题的. 考虑

$$\frac{\partial f}{\partial x}=\begin{cases}2x\dfrac{x^4+y^4}{x^2+y^2}+4x^3\log(x^2+y^2), & x,y\neq 0,\\ 0, & x=y=0.\end{cases}$$

在邻域 $x^2+y^2<\delta^2$ 内, 对于 $\delta<1$, 有

$$\frac{\partial f}{\partial x}<2\delta^3+8\delta^2|\delta\log\delta|<10\delta^2,$$

这里我们已经用了当 $\delta<1$ 时, $|\delta\log\delta|<1$.

练习 1.4 d (p.42)

1. (a) $2ab$,

(c) $abf''(ax+by)$,

(e) $-\dfrac{1}{(x+y)^2}$.

2. (b) $f_x=y\sinh xy$, $f_y=x\sinh xy$, $f_{xx}=y^2\cosh xy$,

$f_{xy}=xy\cosh xy+\sinh xy$, $f_{yy}=x^2\cosh xy$,

$f_{xxx}=y^3\sinh xy$, $f_{xxy}=xy^2\sinh xy+2y\cosh xy$,

$f_{xyy}=x^2y\sinh xy+2x\cosh xy$, $f_{yyy}=x^3\sinh xy$.

(d) $f_x=1/y-y/x^2$, $f_y=1/x-x/y^2$, $f_{xx}=2y/x^3$,

$f_{xy}=(-1/x^2)-1/y^2$, $f_{yy}=2x/y^3$, $f_{xxx}=-6y/x^4$,

$f_{xxy}=2/x^3$, $f_{xyy}=2/y^3$, $f_{yyy}=-6x/y^4$.

问题 1.4 d (p.43)

1. (b) 令 $z=\log u$. 则 $z_{xy}=0$. 这样 z_x 就不依赖于 y. 令 $z_x=\alpha(x)$, 则

$$z=\int\alpha(x)dx+\psi(y)=\phi(x)+\psi(y);$$

从而

$$u=e^z=e^{\phi(x)}\cdot e^{\psi(y)}.$$

练习 1.5 a (p.46)

1. (a), (b) $f_x(0,0)$ 不存在.

(c) 令 $h=\rho\cos\theta$, $k=\rho\sin\theta$. 对于可微性下式成立是必要的:

$$f(h,k)-f(0,0)=\rho\sin 2\theta=f_x(0,0)h+f_y(0,0)k+o(\rho),$$

然而 $f_x(0,0)=f_y(0,0)=0$ 却意味着上式不成立.

2. 对于 x 与 $x+\delta_1$ 之间的 s, y 与 $y+\delta_2$ 之间的 t, 我们有

$$|g(s)-g(x)|<\varepsilon_1(\delta_1), |h(t)-h(y)|<\varepsilon_2(\delta_2),$$

其中 $\lim\limits_{\delta_1\to 0}\varepsilon_1(\delta_1)=\lim\limits_{\delta_2\to 0}\varepsilon_2(\delta_2)=0$.

因此, 用微积分学中的积分中值定理, 我们有

$$\int_{x_0}^{x+\delta_1} g(s)ds=\int_{x_0}^{x} g(s)ds+\delta_1 g(\xi),$$

其中 $|g(\xi)-g(x)|<\varepsilon_1(\delta_1)$; 类似的结果对 $h(t)$ 也成立. 于是推知

$$\begin{aligned}&f(x+\delta_1,y+\delta_2)\\&=\left[\int_{x_0}^{x} g(s)ds+\delta_1 g(x)+o(\delta_1)\right]\left[\int_{y_0}^{y} h(t)dt+\delta_2 h(y)+o(\delta_2)\right]\\&=f(x,y)+\delta_1 g(x)+\delta_2 h(y)+o(\sqrt{\delta_1^2+\delta_2^2}).\end{aligned}$$

问题 1.5 a (p.47)

1. 令 $\rho=\sqrt{h^2+k^2}$. 则

$$|f(x,y)-f(a,b)|\le\rho(|f_x(a,b)|+|f_y(a,b)|+\varepsilon),$$

其中 $\lim\limits_{\rho\to 0}\varepsilon=0$. 于是, f 不仅是连续的, 而且是利普希茨连续: 对于 $P=(x,y)$, $A=(a,b)$, 在 A 的某个邻域我们有

$$|f(P)-f(A)|\le M|P-A|,$$

其中 M 是常数.

练习 1.5 b (p.50)

1. 曲面 $z=f(x,y)$ 与平面 $\arctan[(y-y_0)/(x-x_0)]=\alpha$ 的交线的斜率, 也就是在 $z\rho$ 平面上看, 曲线

$$z=\phi(\rho)=f(x_0+\rho\cos\alpha, y_0+\rho\sin\alpha)$$

的斜率.

2. (a) $a, \dfrac{a\sqrt{3}+b}{2}, \dfrac{a+b\sqrt{3}}{2}, b$;
(c) $2, \sqrt{3}-2, 1-2\sqrt{3}, -4$;
(e) $-1, -\dfrac{\sqrt{3}}{2}, -\dfrac{1}{2}, 0$;
(g) 0,0,0,0.

3. (a) $-8/5$,
(b) -1,
(c) $-2/\sqrt{3}$.

4. $f(x,y)=xy/(x^2+y^2)$.

6. $\partial^2 f/\partial r^2=\sin 2\theta$.

练习 1.5 c (p.53)

1. (a) $z=8y-4$,
(c) $3x+3y-4z+5-3\log 2=0$,
(e) $z=[\exp(1/\sqrt{2})/\sqrt{2}](x-y+\sqrt{2}+\pi/4)$,
(g) $z=2e^{-2}\left(x+y+\dfrac{1}{2}e^2\displaystyle\int_0^2 e^{-t^2}dt-2\right)$.

2. 公共点是原点.

3. 通过三个点的平面方程能写成如下形式:

$$\begin{aligned}z-z_0=&\{(x-x_0)[k_1(z_2-z_0)-k_2(z_1-z_0)]+(y-y_0)\\&\times[h_2(z_1-z_0)-h_1(z_2-z_0)]\}/\{h_2k_1-h_1k_2\},\end{aligned}$$

其中 $h_i = x_i - x_0$, $k_i = y_i - y_0$, $i = 1,2$. 令 $h_i = \rho_i \cos\alpha_i$, $k_i = \rho_i \sin\alpha_i$, 则 $z_i - z_0 = \rho_i[(\cos\alpha_i)(\partial z/\partial x) + (\sin\alpha_i)(\partial z/\partial y) + o(\rho_i)$. 将这些代入平面方程, 则由于 $\sin(\alpha_1 - \alpha_2) \neq 0$, 并且 (x,y) 固定, 便得到需要的结果:

$$z - z_0 = (x - x_0)\frac{\partial z}{\partial x} + (y - y_0)\frac{\partial z}{\partial y} + \frac{o(\rho_2)}{\rho_2} + \frac{o(\rho_1)}{\rho_1}.$$

4. 我们可以假定系数不全为 0, 比如说 $c \neq 0$. 则 (x_0, y_0, z_0) 落在如下两个曲面之一上,

$$z = \pm\sqrt{\frac{1 - ax^2 - by^2}{c}}.$$

切平面的方程是

$$z - z_0 = (x - x_0)z_x(x_0, y_0) + (y - y_0)z_y(x_0, y_0).$$

对二次曲面的方程求微分, 得到

$$2ax_0 + 2cz_0\frac{\partial z}{\partial x} = 0,$$

$$2by_0 + 2cz_0\frac{\partial z}{\partial y} = 0;$$

而把由此求得 $\dfrac{\partial z}{\partial x}$ 和 $\dfrac{\partial z}{\partial y}$ 的值代入切平面方程便得到 (若 $z_0 \neq 0$)

$$z - z_0 = -\frac{ax_0}{cz_0}(x - x_0) - \frac{by_0}{cz_0}(y - y_0);$$

从而

$$ax_0x + by_0y + cz_0z = ax_0^2 + by_0^2 + cz_0^2 = 1.$$

练习 1.5 d (p.56)

1. (a) $(2xy^2 + 3y^3)dx + (2x^2y + 9xy^2 - 8y^3)dy$,

(c) $4x^3dx - 3y^2dy/(x^4 - y^4)$,

(e) $-(dx+y^{-1}dy)\sin(x+\log y)$,

(g) $dy+dy/(1+(x+y)^2)$,

(i) $(dx+dy-dz)\sinh(x+y-z)$.

2. $(-2/10)+(7\sqrt[3]{5}/25)$.

3. $e^{x^2+y^2}[(8x^3+12x)dx^3+(8x^2y+4y)dx^2dy+(8xy^2$
$+4x)dxdy^2+(8y^3+12y)dy^3]$.

练习 1.5 e (p.58)

1. 变量 z 由 -3 变到 -3.5.
2. $-\dfrac{1}{600}$.
3. $1/2(y|h|+x|k|)$.
4. 由 $dz=ydx+xdy$ 有 $dz/z=dx/x+dy/y$.
5. 由 $dg=2dx/t^2-4xdt/t^3$ 知 g 的相对误差是

$$dg/g=dx/x-2dt/t.$$

于是在测量 t 时所产生的一定的相对误差将有在测量 x 时所产生的同一相对误差效果的两倍.

练习 1.6 a (p.63)

1. (a) $z_x=-2x\log(1+y)$, $z_y=-\dfrac{x^2}{1+y}$,

$$z_{xx}=-2\log(1+y), z_{xy}=-\frac{2x}{(1+y)}, z_{yy}=\frac{x^2}{(1+y)^2}.$$

(e) 令 $u=x$, $v=\arctan y$. 则 $z_x=v\sec^2(uv)$, $zy=[\sec^2(uv)]/(1+y^2)$, $z_{xx}=2v^2\sec^2(uv)\tan(uv)$, $z_{xy}=[\sec^2(uv)/(1+y^2)]\times[1+2v\tan(uv)]$, $z_{yy}=x\sec^2(uv)/(1+y^2)^2[x\tan(uv)-2y]$.

2. (a) $w_x=\dfrac{-x-y\cos z}{(x^2+y^2+2xy\cos z)^{3/2}}$,

$$w_y=\frac{-y-x\cos z}{(x^2+y^2+2xy\cos z)^{3/2}},$$

$$w_z = \frac{xy\sin z}{(x^2+y^2+2xy\cos z)^{3/2}};$$

(b) $$w_x = \frac{1}{\sqrt{z^2+2zy^2+y^4-x^2}},$$

$$w_y = \frac{-2xy}{(z+y^2)\sqrt{z^2+2zy^2+y^4-x^2}},$$

$$w_z = \frac{-x}{(z+y^2)\sqrt{z^2+2zy^2+y^4-x^2}}$$

(c) $$w_x = 2x + \frac{2xy}{1+x^2+y^2+z^2},$$

$$w_y = \log(1+x^2+y^2+z^2) + \frac{2y^2}{1+x^2+y^2+z^2},$$

$$w_z = \frac{2yz}{1+x^2+y^2+z^2};$$

(d) $$w_x = \frac{1}{2(1+x+yz)\sqrt{x+yz}},$$

$$w_y = \frac{z}{2(1+x+yz)\sqrt{x+yz}},$$

$$w_z = \frac{y}{2(1+x+yz)\sqrt{x+yz}}.$$

3. (a) 考察 $z=u^v$ 的微商，其中 u 和 v 是 x 的函数:

$$\frac{dz}{dx} = vu^{v-1}\frac{du}{dx} + u^v \log u \cdot \frac{dv}{dx}.$$

对 $u=x$, $v=x$ 使用这个公式便得到

$$\frac{d}{dx}(x^x) = x^x(1+\log x).$$

今再对 $u=x$, $v=x^x$ 使用这个公式就得到

$$\frac{d}{dx}(x^{(x^x)}) = x^{(x^x)}x^x\left[\frac{1}{x}+\log x+(\log x)^2\right]$$

(b) 令 $y=1/x$. 则

$$\frac{dz}{dx}=-\frac{1}{x^2}\frac{dz}{dy}.$$

用 $z=(y^y)^y=u^v$, 其中 $u=y, v=y^2$, 得到

$$\frac{dz}{dy}=y^{(y^2+1)}(1+2\log y)=yz(1+2\log y);$$

从而得到

$$\frac{dz}{dx}=\frac{2\log x-1}{x^{3+1/x^2}}.$$

4. 参看问题第 1 题.

5. 在各种情形中, 利用对不同变量的对称性进行计算:

(a) $f_{xx}=\dfrac{y^2-x^2}{(x^2+y^2)^2}$,

(b) $g_{xx}=\dfrac{2x^2-y^2-z^2}{(x^2+y^2)^2}$,

(c) $h_{xx}=\dfrac{6x^2-2y^2-2z^2-2w^2}{(x^2+y^2)^3}$.

问题 1.6 a (p.64)

1. 在下式中用柯西 - 黎曼方程:

$$\phi_{xx}+\phi_{yy}=(u_x^2+u_y^2)f_{uu}+2(u_xv_x+u_yv_y)f_{uv}$$
$$+(v_x^2+v_y^2)f_{uv}+(u_{xx}+u_{yy})f_u+(v_{xx}+v_{yy})f_v,$$

并且注意 u,v 也是拉普拉斯方程的解.

2. 设圆锥的顶点在原点 (这是不失普遍性的, 因为轴的平移不会影响 f 的微商). 如果点 (x,y,z) 在圆锥上, 那么对一切实数 λ, 点 $(\lambda x,\lambda y,\lambda z)$ 也在圆锥上. 因此, 我们有

$$\frac{z}{x}=f\left(\frac{x}{x},\frac{y}{x}\right)=f\left(1,\frac{y}{x}\right)=\phi\left(\frac{y}{x}\right);$$

这样, 圆锥的方程便能写成一个实变量的函数 ϕ:

$$z=x\cdot\phi\left(\frac{y}{x}\right).$$

两边取微分便得结果.

3. (a) $g_{rr}+\frac{2}{r}g_r$,

(b) 从 $g_{rr}/g_r=-2/r$, 得到 $\log g_r=-2\log r+$ 常数, 等等.

4. (a) $g_{rr}+\frac{n-1}{r}g_r$,

(b) 若 $n=1$, $ar+b$;

若 $n=2$, $a\log r+b$;

若 $n>2$, $a/r^{n-2}+b$(参看问题第 3 题).

练习 1.6 c (p.70)

1. $\sqrt{u_r^2+(1/r^2)u_\theta^2}$.

2. 令 $u=f(x,y)$ 并引进新变量 $\xi=x\cos\theta+y\sin\theta$, $\eta=y\cos\theta-x\sin\theta$, 便得到

$$u_{xx}=\cos^2\theta u_{\xi\xi}-2\cos\theta\sin\theta u_{\xi\eta}+\sin^2\theta u_{\eta\eta},$$

$$u_{yy}=\sin^2\theta u_{\xi\xi}+2\cos\theta\sin\theta u_{\xi\eta}+\cos^2\theta u_{\eta\eta}.$$

4. $z_x=3$, $z_y=1$, $z_r=z_x\cos\theta+z_y\sin\theta$, $z_\theta=-z_xr\sin\theta+z_yr\cos\theta$.

5. 注意微商不依赖于 a 和 b、变换实质上是 x 轴和 y 轴的旋转和平移, 参看练习第 2,3 题. 利用

$$u_{xx}=\alpha^2U_{\xi\xi}-2\alpha\beta U_{\xi\eta}+\beta^2U_{\eta\eta},$$

$$u_{xy}=\alpha\beta U_{\xi\xi}+(\alpha^2-\beta^2)U_{\xi\eta}-\alpha\beta U_{\eta\eta},$$

$$u_{yy}=\beta^2U_{\xi\xi}+2\alpha\beta U_{\xi\eta}+\alpha^2U_{\eta\eta}.$$

其几何意义见问题 1.6a, 第 2 题.

6. $\frac{z^3}{\partial x^2}T_z+T_{xx}+\frac{z}{x}T_{xz}+\frac{z^2}{x^2}T_{zz}$.

问题 1.6 c (p.71)

1. $\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right)+\frac{1}{\sin\theta}\frac{\partial^2 u}{\partial\phi^2}+\frac{\partial}{\partial u}\left(\sin\theta\frac{\partial u}{\partial\theta}\right).$
与问题 1.6a 第 3 题比较，令 u 关于 θ 和 ϕ 的微商等于零.

2. 在给定的变换下，方程

$$Af_{xx}+2Bf_{xy}+Cf_{yy}=0$$

被变换成

$$A^*f_{\xi\xi}+2B^*f_{\xi\eta}+C^*f_{\eta\eta}=0,$$

其中

$$A^*=a^2A+2abB+b^2C.$$
$$B^*=aCA+(ad+bc)B+bdC,$$
$$C^*=c^2A+2cdB+d^2C.$$

(参看练习 1.6c 第 3 题) 注意到

$$B^{*2}-A^*C^*=(ad-bc)^2(B^2-AC),$$

即可见 $B^{*2}-A^*C^*$ 的符号是不随线性变换改变的. 由此推知，下面两种情况是不存在这样线性变换的：对于 (a), 如果

$$B^2-AC\geq 0;$$

对于 (b), 如果 $B^2-AC<0$.

(a) 假设 $B^2-AC<0$, 并设上述 $A^*=1$, $B^*=0$, $C^*=0$. 从 $AC>B^2\geq 0$ 可见 A 和 C 有相同的非零符号，不妨假设它是正的. 若 $B=0$ 取 $b=c=0$, $a=1/\sqrt{A}$, $d=1/\sqrt{C}$. 如果 $B\neq 0$, 可先化为 $B=0$ 的情形，例如，通过取

$$b=0, a=\frac{1}{\sqrt{A}}, C=\frac{B}{\sqrt{A(AC-B^2)}},$$
$$d=\frac{-A}{\sqrt{A(AC-B^2)}}.$$

(b) 假设 $B^2 - AC > 0$ 并设上述 $A^* = C^* = 0$, $B^* = 1$. 如果 $B = 0$, 则 A 和 C 有相反的符号. 在这样情形下满足方程

$$\frac{a}{b} = \sqrt{-\frac{C}{A}}, \frac{d}{c} = \sqrt{-\frac{C}{A}}, bc\sqrt{-AC} = 1;$$

例如, 取

$$a = 1,\ b = \sqrt{-\frac{A}{C}},\ C = \frac{1}{2},\ d = \frac{1}{2}\sqrt{-\frac{c}{A}}.$$

如果 $B \neq 0$ 并且 A 和 C 中至少有一个不为 0, 比如说 $A > 0$, 可先化为 $B = 0$ 的情形, 例如通过取

$$A^* = A, \qquad C^* = -\frac{1}{A}, \quad b = 0,$$

则

$$a = 1,\ d = \frac{1}{\sqrt{B^2 - AC}},\ C = -\frac{B}{\sqrt{A(B^2 - AC)}}.$$

练习 1.7 a (p.73)

1. (a) $(k + h)\cos(x + h + y + k)$,

(b) $-\dfrac{h(y + k)}{(x + h)^2} + \dfrac{k}{x + h}$.

2. (a) $-\dfrac{1}{8}$,

(b) $\dfrac{5}{8}e^{5/16}$,

(c) $\dfrac{\pi}{8}$.

练习 1.7 b (p.75)

1. 通过点 (x, y) 的铅直平面 $h(\eta - y) - k(\xi - x) = 0$ 与曲面 $z = f(x, y)$ 相交, 定义了一条曲线, 在这条曲线的任一段弧上, 存在一个内点, 该点的切线平行于连接两端点的弦.

2. (a) $\dfrac{1}{2}$,

(b) $\dfrac{8}{3\pi}\arcsin\dfrac{8-4\sqrt{2-\sqrt{2}}}{3\pi}$.

3. 取 $x=0, y=-\dfrac{1}{2}, h=k=\dfrac{1}{2}$.

5. (a) $\dfrac{3}{7}$,

(b) $\dfrac{23}{54}$.

问题 1.7 b (p.75)

1. 只需证明，对于任何两点，只要它们能用落在定义域内的线段联接，f 就有同一数值.

练习 1.7 c (p.77)

1. xy.

2. 注意到微分 df 在点 (2,3) 处对于 $h=0.1, k=-0.1$ 等于零，即可见，近似地有

$$f(2.1,2.9)=f(2,3)+\frac{1}{2}d^2f(2,3)=79.9.$$

3. 这个近似是精确的，一切阶的误差都是 0.

4. (a) $x^3-2x^2y+y^2-h(3x^2-4xy)+k(2y-2x^2)$
$\quad +h^2(3x-2y)-kh4x+k^2+6h^3-2h^2k$;

(b) $\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^n(h+2k)^{2n-1}}{(2n-1)!}$;

(c) 必须分两种情形 $x+h>0$ 与 $x+h<0$, 这两种情形产生不同的 h 的一阶项:

$$\begin{aligned}&x^4y-2y^2x-\sqrt{3}|x|+h(4x^3y-2y^2-\sqrt{3}\operatorname{sgn}(x+h))\\&\quad+k(x^4-4yx)+h^26x^2y+4hkx^3-k^22x+h^34xy\\&\quad+h^2k6x^2-2k^2h+h^4y+4h^3k+h^4k.\end{aligned}$$

5. $x+x(y-1)-2x(z+1)-2x(y-1)(z+1)$
$\quad +2x(z+1)^2+x(y-1)(z+1)^2$.

6. (a) $y - x^2 - \frac{y^3}{3} + x^4y - x^2y^3 + \frac{y^5}{5} + \cdots$,

(b) $y + \frac{x^2y}{2} + \frac{y^3}{6} + \frac{x^2y^3}{12} + \frac{x^4y}{24} + \frac{x^5}{120} + \cdots$,

(c) $1 + y + \frac{y^2}{2} - \frac{x^4}{6} - \frac{x^3y}{3} + \frac{xy^3}{6} + \frac{y^4}{24} + \cdots$,

(d) $1 + x + \frac{x^2}{2} - \frac{y^2}{2} + \frac{x^3}{6} - \frac{xy^2}{2} + \cdots$,

(e) $x - \frac{x^3}{6} + \frac{xy^2}{2} + \frac{x^5}{120} - \frac{x^3y^2}{12} + \frac{5xy^4}{24} + \cdots$,

(f) $xy + \frac{x^2y}{2} + \frac{xy^2}{2} + \frac{x^3y}{3} + \frac{x^2y^2}{4} + \frac{xy^3}{3} + \cdots$,

(g) $1 + x^2 - y^2 + \frac{x^4}{2} - x^2y^2 + \frac{y^4}{2} + \cdots$,

(h) $1 - \frac{3}{2}x^2 - xy - \frac{y^2}{2} + \cdots$,

(i) $1 - \frac{x^2}{2} - \frac{x^2y^2}{2} + \frac{x^4}{24} - \frac{x^6}{120} - \frac{x^4y^2}{12} + \frac{x^2y^4}{3} + \cdots$,

(j) $x^2 + y^2 - \frac{x^6}{6} - \frac{x^4y^2}{2} - \frac{x^2y^4}{2} - \frac{y^6}{6} + \cdots$.

7. 注意误差是 4 阶的，展开到四阶有

$$\frac{\cos x}{\cos y} = 1 - \frac{x^2 - y^2}{2} + \frac{x^4 - 6x^2y^2 + 5y^4}{24} + \cdots;$$

对于这 4 阶项我们有

$$\frac{x^4 - 6x^2y^2 + 5y^4}{24} = \frac{(y^2 - x^2)(5y^2 - x^2)}{24}.$$

对于 $|x| \le \pi/6$, $|y| \le \pi/6$ 这两个因子在 $x = y = \frac{\pi}{6}$ 时达到它们的极大值. 于是，我们估计到误差大约是

$$\frac{5}{24}\left(\frac{\pi}{6}\right)^4 \approx 0.016.$$

问题 1.7 c (p.78)

$$1.\ \text{(a)}\ \sum_{n=0}^{\infty}\sum_{r=0}^{n}\binom{n}{r}x^r y^{n-r}=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\binom{m+n}{n}x^m y^n;$$

在带形域 $|x+y|<1$ 内收敛.

$$\text{(b)}\ \sum_{n=0}^{\infty}\sum_{r=1}^{n}\frac{x^r}{r!}\frac{y^{n-r}}{(n-r)!}=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{x^m}{m!}\cdot\frac{y^n}{n!},$$

对 x 和 y 的一切值收敛.

2. 展开球面公式的两边直到 x, y 和 z 的二阶.

3. 在 $(h, e^{-1/h})$ 的邻域内展开 $f(2h, e^{-1/2h})$ 与 $f(0,0)$ 到二阶; 求和并除以 h^2.

4. 收敛性可从单变量指数函数展开式的收敛性直接推得. 对 x 求微商得到

$$2yf(x,y)=\sum_{n=0}^{\infty}\frac{H_n'(x)y^n}{n!}=\sum_{n=1}^{\infty}\frac{2H_{n-1}(x)y^n}{(n-1)!};$$

从而令相应系数相等即得 (b). 从 (b) 和 $H_0(x)=1$ 可用数学归纳法得到 (a). 为了得到 (c), 对 y 求微商并且让相应系数相等. 为了得到 (d), 用 (b), 在 (c) 中用 H_n' 代替 $2nH_{n-1}$, 然后求微商得到

$$H_{n-1}'-2xH_n'+2H_n'+H_n''=0.$$

最后, 在这些结果中用 (b) 把 H_{n+1}' 换成 $2(n+1)H_n$.

练习 1.8 b (p.88)

1. 在闭区间 $a\le x\le b$ 上用 $\beta_k(x,k)$ 对 x 的一致连续性并限制 k 在区间 $k_0<k<k_1$ 内部的任一闭区间上.

2. (a) 对于 $\epsilon=k^{-2/3}$ 与 $1-\epsilon<x<1$, 当 k 很大时我们有

$$k\log x=k(x-1)+O(k^{-1/3}),$$

$$\frac{x-1}{\log x}=1+O(k^{-2/3}),$$

因此

$$\frac{x^k(x-1)}{\log x}=e^{k(x-1)}(1+O(k^{-1/3}));$$

而对于 $0<x<1-\varepsilon$ 我们有

$$\frac{x^k(x-1)}{\log x}=O\Big(\frac{x-1}{\log x}e^{-k^{1/3}}\Big).$$

由此推知

$$F(k)=\int_{1-\varepsilon}^{1}+\int_{0}^{1-\varepsilon}=\frac{1}{k}+O(k^{-4/3}).$$

(b) 按照练习题 1 有

$$F'(k)=\int_0^1 x^k(x-1)dx=\frac{1}{k+2}-\frac{1}{k+1}.$$

因此,

$$F(k)=\log\frac{2+k}{1+k}+c,$$

而根据 (a), 这里常数 (c) 原来是 0.

练习 1.9 b (p.100)

1. (a) $\displaystyle\int_0^{2\pi}(-t\sin t+\cos^2 t+\sin t)dt=3\pi,$

(b) $$\int_{-1}^{+1}(-2t^2x_0-2tx_0y_0(1-t^2)+y_0(1-t^2))dt$$

$$=-\frac{4}{3}(x_0-y_0).$$

第二章

练习 2.1 (p.152)

1. 如果 $X=(x,y,z)$ 是直线上任意一点，那么

$$\overrightarrow{PX}=\lambda\mathbf{A},$$

其中 λ 可以是任何实数. 于是

$$(x+2,y,z-4)=\lambda(2,1,3),$$

或

$$\frac{x+2}{2}=y=\frac{z-4}{3}.$$

2. 设 $\overrightarrow{PQ}=\mathbf{A}$. 直线上任何一点 X 满足 $\overrightarrow{PX}=\lambda\mathbf{A}$. 设 $\mathbf{B},\mathbf{C}$ 和 $\mathbf{V}$ 分别是 P,Q 和 X 的位置向量，则

$$\overrightarrow{PX}=\mathbf{V}-\mathbf{B}=\lambda\mathbf{A}=\lambda(\mathbf{C}-\mathbf{B});$$

或

$$\mathbf{V}=(1-\lambda)\mathbf{B}+\lambda\mathbf{C}.$$

特别是，若 $P=(3,-2,2)$ 和 $Q=(6,-5,4)$, 如在 (a) 中所给定的，

$$(x,y,z)=\lambda(3,-3,2),$$

或

$$\frac{x}{3}=-\frac{y}{3}=\frac{z}{2}.$$

3. 如果 $\mathbf{V}$ 是联结 P 到 Q 的线段上任何一点 X 的位置向量，那么根据练习第 2 题的解答，对于某实数 λ 有

$$\mathbf{V}=(1-\lambda)\mathbf{A}+\lambda\mathbf{B}.$$

于是

$$(1-\lambda)(\mathbf{V}-\mathbf{A})=\lambda(\mathbf{B}-\mathbf{V})=(1-\lambda)\lambda(\mathbf{B}-\mathbf{A}).$$

如果 $0<\lambda<1$, 则有 $\mathbf{V}-\mathbf{A}$, $\mathbf{B}-\mathbf{V}$ 与 $\mathbf{B}-\mathbf{A}$ 方向相同并且

$$|\mathbf{V}-\mathbf{A}|/|\mathbf{B}-\mathbf{V}|=\lambda/(1-\lambda).$$

4. 将位置向量写成如下形式

$$\mathbf{V}=\mathbf{A}+\lambda(\mathbf{B}-\mathbf{A}),$$

其中 $\mathbf{B}-\mathbf{A}$ 是由 $\overrightarrow{PQ}$ 表示的，可见 $\lambda>0$.

5. 设 $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{E}$ 分别是点 P,Q,R,S,M 的位置向量. 按 1/3 的比例分 MS 于 O 点，并取 O 点为原点. 于是，

$$\mathbf{D}=-3\mathbf{E}.$$

因为

$$\mathbf{E}=1/3(\mathbf{A}+\mathbf{B}+\mathbf{C}),$$

所以

$$\frac{1}{4}(\mathbf{A}+\mathbf{B}+\mathbf{C}+\mathbf{D})=\mathbf{O}.$$

因此，按一般定义， O 是质量中心，因而它的位置显然与顶点的顺序无关.

6. 设这两条边是 PQ 和 RS; 用前题解答的符号，它们的中点分别具有位置向量

$$\frac{1}{2}(\mathbf{A}+\mathbf{B})\ 和\ \frac{1}{2}(\mathbf{C}+\mathbf{D}).$$

从练习第 5 题的解答知道

$$\frac{1}{2}(\mathbf{A}+\mathbf{B})=-\frac{1}{2}(\mathbf{C}+\mathbf{D});$$

因此，两个中点与质心 O 共线并且到 O 等距.

7. 如果 $P_k=(x_k,y_k,z_k)$, $k=1,2,\cdots n$, 则

$$G=(x_0,y_0,z_0)=\left(\frac{\Sigma m_kx_k}{\Sigma m_k},\frac{\Sigma m_Ky_k}{\Sigma m_k},\frac{\Sigma m_kz_k}{\Sigma m_k}\right),$$

$$\Sigma m_kA_K=(\Sigma m_K(x_K-x_0),\ \Sigma m_K(y_K-y_0),\Sigma m_K(z_K-z_0))=(0,0,0)$$

8. 零向量是实数 1. “向量”a 与数 λ 的“乘法”意味着将 a 升 λ 次幂. 于是，如果向量“加法”用 $\oplus$ 表示，数乘用 $\odot$ 表示，则

$$\lambda\odot(a\oplus b)=(ab)^\lambda=a^\lambda\cdot b^\lambda=(\lambda\odot a)\oplus(\lambda\odot b).$$

9. 复数 $a+ib$ 对应到向量 (a,b).

10. 取原点为球的中心并设 $\mathbf{A},\mathbf{B},\mathbf{R}$ 分别是 P,Q,R 的位置向量. 如果球的半径为 ρ, 则

$$|\mathbf{A}|^2=|\mathbf{B}|^2=|\mathbf{R}|^2=\rho^2$$

并且 $\mathbf{B}=-\mathbf{A}$. 由此, 根据 (15c), 就得到

$$(\mathbf{R}-\mathbf{A})\cdot(\mathbf{R}-\mathbf{B})=(\mathbf{R}-\mathbf{A})\cdot(\mathbf{R}+\mathbf{A})=|\mathbf{R}|^2-|\mathbf{A}|^2=0.$$

11. (a) 根据 $(\mathbf{X}-\mathbf{P})\cdot\mathbf{A}=0$, 平面的方程是

$$x+2y-2z=-1.$$

由于单位法向量 $\mathbf{B}=(-1/3,-2/3,2/3)$, 故得法线方程

$$-\frac{1}{3}x-\frac{2}{3}y+\frac{2}{3}z=\frac{1}{3}.$$

(b) 2/3.

(c) 一样.

12. (a) 令 $P=(y_1,y_2,\cdots,y_n)$ 并设 $\mathbf{B}$ 是 P 的位置向量. 若 $Q=(x_1,x_2,\cdots x_n)$ 是垂足, 它的位置向量为 $\mathbf{X}$, 则

$$\mathbf{A}\cdot\mathbf{X}=c \text{ 和 } \mathbf{B}-\mathbf{X}=\lambda\mathbf{A}.$$

于是 $\mathbf{A}\cdot(\mathbf{B}-\lambda\mathbf{A})=c$, 因此 $\lambda=(\mathbf{A}\cdot\mathbf{B}-c)/|\mathbf{A}|^2$ 而且

$$\mathbf{X}=\mathbf{B}+\mathbf{A}(c-\mathbf{A}\cdot\mathbf{B})/|\mathbf{A}|^2.$$

(b) 分别是 $(-1/9,2/9,2/9)$ 和 $(7/9,-13/9,-5/9)$.

13. 首先注意 $\mathbf{C}\neq\mathbf{O}$; 因为否则就会有

$$\mathbf{A}=\frac{\mathbf{A}\cdot\mathbf{B}}{|\mathbf{B}|^2}\mathbf{B},$$

这与 $\mathbf{A}$ 与 $\mathbf{B}$ 不相平行的条件矛盾. 于是有 $\mathbf{B}\cdot\mathbf{C}=0$.

14. 直线与平面的夹角是直线与平面法线的夹角的余角；所以

$$\sin\phi=\frac{\alpha A+\beta B+\gamma C}{\sqrt{\alpha^2+\beta^2+\gamma^2}\sqrt{A^2+B^2+C^2}}.$$

练习 2.2 (p.171)

1. (a) 直线 $x=-1+4\lambda,\ y=2,\ z=1+3\lambda$;

(b) 平面 $x=2+3\mu+\nu,\ y=1-2\mu,$

$z=-4+\mu-\nu$; 或 $x+2y+z=0$;

(c) 满足 $x+2y+z=0$ 与 $2y+2z+W=-4$ 的点 (x,y,z,w) 所形成的二维线性空间.

2. (a). $\mathbf{A}_1=\sqrt{2}\mathbf{E}_1+2\mathbf{E}_3$.

3. 对于 $\mathbf{E}_1$, 只有 $\mathbf{E}_1=\mathbf{A}_1/|\mathbf{A}_1|$ 是可能的. 假设这样的向量一直到指标 $k-1$ 已经被找到了. 取 $\mathbf{E}_k=\mathbf{V}_k/|\mathbf{V}_K|$, 其中

$$\mathbf{V}_k=\mathbf{A}_k-\sum_{\mu=1}^{k-1}(\mathbf{A}_\mu\cdot\mathbf{E}_\mu)\mathbf{E}_\mu.$$

注意, 如果 $\mathbf{E}_\mu$ 依赖于 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_\mu(\mu=1,2,\cdots k-1)$, 则 $\mathbf{E}_k$ 依赖于 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_k$.

4. 设 $\mathbf{A}_k(k=1,2,\cdots n+1)$ 是任意 $n+1$ 个向量的集合, 如果 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n$ 是相关的, 则 $n+1$ 个向量的一个集合也是相关的; 如果不是这样, 就根据练习第 3 题知道, 向量 $\mathbf{E}_1,\cdots,\mathbf{E}_n$ 依赖于 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n$, 因为 $\mathbf{E}_k(k=1,2,\cdots,n)$ 可以取作坐标向量, 所以 $\mathbf{A}_{n+1}$ 依赖于 $\mathbf{E}_1,\mathbf{E}_2,\cdots,\mathbf{E}_n$; 因此, 进一步它依赖于 $\mathbf{A}_1,\mathbf{A}_2,\cdots,\mathbf{A}_n$.

5. 用向量形式写出直线方程

$$\mathbf{Z}=\mathbf{A}\tau+\mathbf{B},$$

其中 $\mathbf{B}=(b,d,f)$, $\mathbf{A}=(a,c,e)$. 设 Q 是从 P 到该直线的垂足并且 $\mathbf{X}_0=(x_0,y_0,z_0)$, $\mathbf{X}_1=(x_1,y_1,z_1)$ 分别是 P,Q 的位置向量. 因为 Q

在直线上, 对于某个数 τ 有 $\mathbf{X}_1 = \mathbf{A}\tau+\mathbf{B}$. 此外, 从 $(\mathbf{X}_1-\mathbf{X}_0)\cdot\mathbf{A} = 0$ 可知所求的距离 d 由下式给定:

$$d^2 = |\mathbf{X}_1 - \mathbf{X}_0|^2 = (\mathbf{X}_1 - \mathbf{X}_0) \cdot (\mathbf{A}\tau + \mathbf{B} - \mathbf{X}_0)$$

$$= (\mathbf{X}_1 - \mathbf{X}_0) \cdot (\mathbf{B} - \mathbf{X}_0),$$

$$= (x_1 - x_0)(b - x_0) + (y_1 - y_0)(d - y_0)$$

$$+ (z_1 - z_0)(f - z_0),$$

其中

$$(x_1, y_1, z_1) = (a\tau + b, c\tau + d, e\tau + f),$$

$$\tau = \frac{(\mathbf{X}_0 - \mathbf{B}) \cdot \mathbf{A}}{|\mathbf{A}|^2}$$

$$= \frac{a(x_0 + b) + c(y_0 + d) + e(z_0 - f)}{a^2 + c^2 + e^2}.$$

6. 不是. 为了证明这个结论, 只要证明系数向量 (1,2,3), (2,3,1), (3,1,2) 是线性无关的. 例如, 可用练习第 3 题解答的方法, 构造一向量集合, 它由依赖于系数向量而且互相垂直的向量组成.

7. 这等价于在练习第 6 题中解线性方程组, 在右边以常数 a_1, a_2, a_3 代替 0,0,0

$$x_1 = \frac{1}{18}(-5a_1 + a_2 + 7a_3),$$

$$x_2 = \frac{1}{18}(a_1 + 7a_2 - 5a_3),$$

$$x_3 = \frac{1}{18}(7a_1 - 5a_2 + a_3).$$

8. 根据练习第 7 题的解便得到

$$\frac{1}{18}\begin{pmatrix} -5 & 1 & 7 \\ 1 & 7 & -5 \\ 7 & -5 & 1 \end{pmatrix}.$$

9. 若 $\mathbf{a}$ 是奇异的，则列向量 $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ 是相关的．如果解 $\mathbf{X} = (x_1, x_2, \cdots, x_n)$ 对每个 $\mathbf{Y}$ 存在，那么每个 $\mathbf{Y}$ 都会有表达式

$$\mathbf{Y} = x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n,$$

但是 $\mathbf{A}_K$ 不张成整个空间．

10. $\mathbf{ab} = \begin{pmatrix} -2, & 3, & 4 \\ 1 & 0 & 1 \\ -4 & 3 & 2 \end{pmatrix}$, $\mathbf{ba} = \begin{pmatrix} -2 & -4 & 1 \\ -4 & -2 & 1 \\ 3 & 3 & 0 \end{pmatrix}$.

11. $\Delta = ad - bc \neq 0$,

$$\frac{1}{\Delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

12. 设 $\mathbf{ae} = \mathbf{ea} = \mathbf{a}$ 并且 $\mathbf{a}'\mathbf{e} = \mathbf{e}'\mathbf{a} = \mathbf{a}$ 对一切方阵 $\mathbf{a}$ 成立，那么

$$\mathbf{e}'\mathbf{e} = \mathbf{ee}' = \mathbf{e} = \mathbf{e}'.$$

13. $\mathbf{b}^{-1}\mathbf{a}^{-1}$.

14. 根据我们的定义，一个矩阵是奇异的，当且仅当列向量是相关的．于是至少有一个列向量可以表示为其余列向量的线性组合由此推知，映象中的任何一个像向量都能表示成不超过 $n-1$ 个给定向量的线性组合．反之，如果象空间的维数小于 n，则矩阵的列向量定是线性相关的；因为假使它们线性无关，它们的线性组合就会张成 n 维空间．

15. 把 $\mathbf{X}$ 表示成 $(\gamma\cos\theta, \gamma\sin\theta)$ 这个形式，则对于

$$\mathbf{a} = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}$$

有

$$a\mathbf{X} = (\gamma\cos(\theta+\gamma), \gamma\sin(\theta+\gamma));$$

因此，$\mathbf{a}$ 可以解释为向量经过角度 γ 的旋转或坐标轴经过角度 $-\gamma$ 的旋转．对于

$$b = \begin{pmatrix} \cos\gamma & \sin\gamma \\ \sin\gamma & -\cos\gamma \end{pmatrix}$$

有

$$bx = (\gamma\cos(\gamma-\theta),\ \gamma\sin(\gamma-\theta));$$

这可看作向量对与 x 轴倾角为 $(1/2)\gamma$ 的直线的反射，或先颠倒 y 轴方向接着整个坐标轴经过角度 $-\gamma$ 的旋转.

16. 根据 (49a), 对于正交性这个条件是必要的. 同时它也是充分的; 因为如果 $\mathbf{A}_k$ 是 $\mathbf{a}$ 的第 k 个列向量，则它就是 $\mathbf{a}^T$ 的第 k 个行向量. 按照矩阵乘法的定义， $\mathbf{aa}^T = e$ 意味着

$$\mathbf{A}_j \cdot \mathbf{A}_k = \begin{cases} 0, & 当\ j \neq k, \\ 1, & 当\ j = k. \end{cases}$$

17. 令 $\mathbf{c} = \mathbf{ab}$, 若 $\mathbf{c} = (c_{ij})$, 则 $\mathbf{c}^T = (c_{ij}^T)$, 其中

$$c_{ij}^T = c_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki} = \sum_{k=1}^{n} b_{ik}^T a_{kj}^T = (\mathbf{b}^T\mathbf{a}^T)_{ij}.$$

18. 根据练习第 13,17 和 16 题，如果 $\mathbf{a}$ 与 $\mathbf{b}$ 是正交的，则

$$(\mathbf{ab})^T = \mathbf{b}^T\mathbf{a}^T = \mathbf{b}^{-1}\mathbf{a}^{-1} = (\mathbf{ab})^{-1}.$$

对于 $\mathbf{ab}$ 的正交性上述等式也是一个充分条件.

19. 若 $\mathbf{X} = (x_1, x_2, \cdots, x_n)$ 和 $\mathbf{Y} = (y_1, y_2, \cdots, y_n)$, 则根据 (47) 有

$$\begin{aligned}(\mathbf{aX})\cdot(\mathbf{aY}) &= (x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + \cdots + x_n\mathbf{A}_n) \\ &\quad \cdot (y_1\mathbf{A}_1 + y_2\mathbf{A}_2 + \cdots + y_n\mathbf{A}_n) \\ &= x_1y_1 + x_2y_2 + \cdots + x_ny_n.\end{aligned}$$

20. 一个保长的矩阵 $\mathbf{a}$ 必定也是保数量积的; 因为

$$\begin{aligned}|\mathbf{aX} + \mathbf{aY}|^2 &= |\mathbf{aX}|^2 + |\mathbf{aY}|^2 + 2(\mathbf{aX})\cdot(\mathbf{aY}) \\ &= |\mathbf{X}|^2 + |\mathbf{Y}|^2 + 2(\mathbf{ax})\cdot(\mathbf{aY})\end{aligned}$$

$$= |\mathbf{a}(\mathbf{X}+\mathbf{Y})|^2 = |\mathbf{X}+\mathbf{Y}|^2$$

$$= |\mathbf{X}|^2 + |\mathbf{Y}|^2 + 2|\mathbf{X}||\mathbf{Y}|$$

(参看练习第 18 题答案). 由此得条件 (47), 因为每个坐标向量 $\mathbf{E}_k$ 被映象到 $\mathbf{a}$ 的列向量 $\mathbf{A}_k$.

21. 设这些质点是 $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_k$, 它们的质量分别是 $m_1, m_2, \cdots, m_k$, 假定进行仿射变换 $\mathbf{X}' = \mathbf{a}\mathbf{X} + \mathbf{A}$. 假设质心在变换前后分别是

$$\mathbf{X}_0 = \Big(\sum_{j=1}^{k} m_j \mathbf{X}_j\Big) \Big/ \sum_{j=1}^{k} m_j, \ \mathbf{Y}_0 = \Big(\sum_{j=1}^{k} m_j \mathbf{X}_j'\Big) \Big/ \sum_{j=1}^{m} m_j,$$

注意 $\mathbf{X}_0' = \mathbf{a}\mathbf{X}_0 + \mathbf{A} = \mathbf{Y}_0$.

练习 2.3 (p.191)

1. (a) 0,

(b) 2,

(c) 12,

(d) $(x-y)(y-z)(z-x)(x+y+z)$.

2. $a + c = 2b$.

3. (a) 应用 $\det(\mathbf{ea}) = \det(\mathbf{a})$,

(b) 应用 $\det(\mathbf{e}) = \det(\mathbf{a}\mathbf{a}^{-1})$.

4. (a) -1,

(b) 1,

(c) -1,

(d) 1.

5. 如果行列式的所有元素都等于 0, 则结果立即可得. 否则, 我们可以假定 $a_{11} \neq 0$; 因为如果 $a_{ij} \neq 0$, 我们可以调换第一行与第 i 行以及第 1 列与第 j 列, 使 a_{ij} 位于第 1 行与第 1 列, 这样做也许会改变行列式的符号. 从第 j 列中减去用 a_{1j}/a_{11} 乘了以后的第 1 列, 使得第 j 列的每一个元素等于 0, 可用类似的方法使任何一

行的每一个元素等于 0. 通过这种运算并且如果必要的话用 -1 乘第 1 行，就可把行列式转变成如下形式

$$\begin{vmatrix} \alpha & 0 & 0 \\ 0 & b_{11} & b_{12} \\ 0 & b_{21} & b_{22} \end{vmatrix}.$$

这个过程同样适用于子行列式

$$\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix},$$

把它转变成

$$\begin{vmatrix} \beta & 0 \\ 0 & \gamma \end{vmatrix}$$

这个形式. 因为在子行列式上进行的运算能够扩张到初始行列式的行和列而不影响第 1 行与第 1 列的零元素，这样就得到了所求的形式.

6. 在 (66a) 中仅可能有的非零项是出现在 $j_1 = 1, j_2 = 2, \cdots, j_n = n$.

7. 在 $a_{j_1 1} a_{j_2 2} \cdots a_{j_n n}$ 中，设 k 是使 $j_k \neq k$ 的一切指标中的最小者. 如果 $j_k < k$, 则乘积等于 0. 如果 $j_k > k$ 则对于某一个因子 a_{km} (其中 $k < m$), k 必定作为一个行指标出现; 因此，这乘积又等于 0, 于是 $a_{11}a_{22}, \cdots, a_{nn}$ 是 (66a) 中仅可能有的非零项.

8. (a) $(x-y)(y-z)(z-x)$,

(b) -12,

(c) $(2!)^2(3!)^2(4!)$.

9. $x=3, y=2, z=1$.

10. 应用 $\det(\mathbf{a}) \cdot \det(\mathbf{b}) = \det(\mathbf{a}^{\mathrm{T}}\mathbf{b})$.

11. 用 $D = (A+2B)(A-B)^2$

$$= [(x+y+z)(x^2+y^2+z^2-xy-yz-xz)]^2.$$

12. 因为这个行列式在列向量上是交错形式的，直接可以得到 $\Delta = A + Bx$. 对于 $x = -a$, 矩阵是下三角形的；对于 $x = -b$, 矩阵是上三角形的. 因此，根据练习第 7 题有

$$A + Ba = f(a) \quad 且\ A + Bb = f(b).$$

13. 根据 (57a), 在 $\mathbf{c} = (c_{jk})$ 的情形有

$$\begin{aligned} f(\mathbf{A}, \mathbf{B}) &= \sum_{j,k=1}^{n} c_{jk} a_j b_k \\ &= \sum_{j=1}^{n} a_j \sum_{k=1}^{n} c_{jk} b_k \\ &= \mathbf{A} \cdot (\mathbf{CB}) \\ &= \sum_{k=1}^{n} b_k \sum_{j=1}^{n} c_{jk} a_j \\ &= \mathbf{B} \cdot (\mathbf{C}^{\mathrm{T}} \mathbf{A}). \end{aligned}$$

14. 令 $\mathbf{X} = (x, y, z)$, $\mathbf{A} = (g, h, i)$, 并且

$$\mathbf{a} = \begin{pmatrix} a & \frac{1}{2}d, & \frac{1}{2}e \\ \frac{1}{2}d & b & \frac{1}{2}f \\ \frac{1}{2}e & \frac{1}{2}f & c \end{pmatrix},$$

再把二次方程改写成如下形式:

$$\mathbf{X} \cdot (\mathbf{aX}) + \mathbf{A} \cdot \mathbf{X} + \mathbf{j} = 0.$$

如果仿射变换给定成这样的形式

$$\mathbf{X}' = \mathbf{bX} + \mathbf{B},$$

它的逆变换就是

$$\mathbf{X} = \mathbf{C}\mathbf{X}' + \mathbf{C},$$

这里 $\mathbf{c} = \mathbf{b}^{-1}$, $\mathbf{C} = -b^{-1}\mathbf{B}$. 于是这二次方程在新坐标系下是

$$\mathbf{C}\mathbf{X}' \cdot (\mathbf{ac}\mathbf{X}') + \mathbf{C}(\mathbf{ac}\mathbf{X}') + \mathbf{C}\mathbf{X}' \cdot (\mathbf{aB})$$

$$+ \mathbf{A} \cdot \mathbf{C}\mathbf{X}' + \mathbf{C} \cdot (\mathbf{ac}) + \mathbf{A} \cdot \mathbf{B} + \mathbf{j} = 0.$$

应用前一练习题的结果便可把这表示成如下形式

$$X' \cdot (\mathbf{a}'\mathbf{X}') + \mathbf{A}' \cdot \mathbf{X}' + \mathbf{j}' = 0,$$

其中

$$\mathbf{a}' = \mathbf{c}^T\mathbf{ac},$$

$$\mathbf{A}' = \mathbf{c}^T(\mathbf{a}^T\mathbf{c} + \mathbf{aB} + \mathbf{A}),$$

$$\mathbf{j}' = \mathbf{C} \cdot \mathbf{ac} + \mathbf{A} \cdot \mathbf{B} + \mathbf{j}.$$

15. 比较齐次线性方程组

$$\begin{cases} a_1x + a_2y + dz = 0, \\ b_1x + b_2y + ez = 0, \\ c_1x + c_2y + fz = 0. \end{cases}$$

如果这个方程组在 $z = -1$ 时有解, 因而是一个非平凡解, 于是行列式 D 必定等于零. 反过来, 如果行列式等于零, 列向量便是相关的. 于是, 存在常数 x, y, z, 不全为 0, 使得

$$x\mathbf{A}_1 + y\mathbf{A}_2 + z\mathbf{B} = 0,$$

其中 $\mathbf{A}_i = (a_i, b_i, c_i)$, $\mathbf{B} = (d, e, f)$. 不可能 $z = 0$, 因为否则 $\mathbf{A}_1$ 与 $\mathbf{A}_2$ 就会是相关的, 因而所有三个 2×2 行列式都会等于 0. 因此, 我们可以用 $-z$ 去除两边使得 $\mathbf{B}$ 的系数为 -1. 可见所要求的解是存在的.

16. 直线可以写成向量形式

$$\mathbf{X} = \mathbf{A}t + \mathbf{B}, \quad \mathbf{X} = \mathbf{C}t + \mathbf{D}.$$

当且仅当 $\mathbf{A}$ 与 $\mathbf{C}$ 相平行时，这两条直线相平行 (这包括两条直线重合的情形). 当且仅当存在两个数 t_1, t_2 使得

$$\mathbf{A}t_1 + \mathbf{B} = \mathbf{C}t_2 + \mathbf{D}$$

时，该两直线相交. 于是，按照前一题的解答，所求条件是，以

$$\mathbf{A}, \mathbf{C}, \mathbf{B} - \mathbf{D}$$

为列向量的矩阵具有等于零的行列式，也就是

$$\begin{vmatrix} a_1 & c_1 & b_1 - d_1 \\ a_2 & c_2 & b_2 - d_2 \\ a_3 & c_3 & b_3 - d_3 \end{vmatrix} = 0.$$

17. 把 $j_1, j_2, \cdots j_n$ 排列成 $1, 2, \cdots, n$ 的一组对调也把 $1, 2, \cdots, n$ 排列成 $k_1, k_2, \cdots, k_n$. 因此， $j_1, j_2, \cdots, j_n$ 与 $k_1, k_2, \cdots, k_n$ 或者都是奇的，或者都是偶的.

18. 这在向量形式下说的是，向量方程

$$\mathbf{a}\mathbf{X} = \lambda\mathbf{X}$$

必定至少有一个非零解. 试把这方程重新改写成齐次方程的形式:

$$(\mathbf{a} - \lambda\mathbf{e})\mathbf{X} = 0,$$

其中 $\mathbf{e}$ 是单位方阵. 这个方程有非平凡解，当且仅当

$$\det\,(\mathbf{a} - \lambda\mathbf{e}) = 0.$$

在 n 维空间中这是 λ 的 n 次多项式方程，具有首项 $(-1)^n\lambda^n$. 因此，当 n 是奇数时，解总是存在的.

练习 2.4 (p.218)

1. 设 $\mathbf{X}_0$ 是点 P 的位置向量，并且把直线 l 表示成向量形式 $\mathbf{X}=\mathbf{A}t+\mathbf{B}$. 从 P 到 l 的距离 r 是 $|\mathbf{X}_0-\mathbf{B}|\sin\theta$, 其中 θ 是 $\mathbf{P}-\mathbf{B}$ 与 $\mathbf{A}$ 之间的夹角. 因此,

$$r=|(\mathbf{X}_0-\mathbf{B})\times\mathbf{A}|/|\mathbf{A}|.$$

2. 速度是 $r\omega$, 这里 r 是点到转轴的距离. 根据前一题的解答, 若以 $\mathbf{B}$ 表示原点 $\mathbf{X}_0=(x,y,z)$ 并且 $\mathbf{A}=(\alpha,\beta,\gamma)$, 则

$$r\omega=\omega[(y\gamma-z\beta)^2+(z\alpha-x\gamma)^2+(x\beta-y\alpha)^2]^{1/2}.$$

3. 把这三个点的位置向量分别叫做 $\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3$. 如果 $\mathbf{X}=(x,y,z)$ 表示平面上任意一个点, 这三个向量 $\mathbf{X}_1-\mathbf{X},\mathbf{X}_2-\mathbf{X},\mathbf{X}_3-\mathbf{X}$ 就落在一个二维空间上, 因而是相关的. 所以

$$\det\,(\mathbf{X}_1-\mathbf{X},\mathbf{X}_2-\mathbf{X},\mathbf{X}_3-\mathbf{X})=0.$$

4. 设该二直线的方程给成了向量形式如 l: $\mathbf{X}=\mathbf{A}t+\mathbf{B}$ 和 $l':\mathbf{X}'=\mathbf{A}'t'+\mathbf{B}'$. 在每条直线上, 各有一个端点的最短线段 $\overline{PP'}$ 必然垂直于这两条直线. 因为, 比如说 $\overline{PP'}$ 不是在 P' 点垂直于 l'; 那么从 P 到 l' 的垂线就会比 $\overline{PP'}$ 短. 如果 $\mathbf{X}$ 和 $\mathbf{X}'$ 分别是 P 和 P' 的位置向量, 则

$$\mathbf{X}-\mathbf{X}'=\mathbf{A}t+\mathbf{B}-\mathbf{A}'t'-\mathbf{B}'=k(\mathbf{A}\times\mathbf{A}').$$

为了决定 k, 对这个方程用 $(\mathbf{A}\times\mathbf{A}')$ 做点乘, 就得出

$$k=\frac{(\mathbf{B}-\mathbf{B}')\cdot(\mathbf{A}\times\mathbf{A}')}{|\mathbf{A}\times\mathbf{A}'|^2},$$

这就给出所要求的距离 d:

$$d^2=|\mathbf{X}-\mathbf{X}'|^2=k^2|\mathbf{A}\times\mathbf{A}'|^2,$$

或

$$d=\frac{|(\mathbf{B}-\mathbf{B}')\cdot(\mathbf{A}\times\mathbf{A}')|}{|\mathbf{A}\times\mathbf{A}'|}.$$

5. 这个和并不依赖于原点的选择，因为选择一个不同的原点 (a,b) 相当于把每个行列式

$$\Delta_k=\begin{vmatrix} x_k & x_{k+1} \\ y_k & y_{k+1} \end{vmatrix} \text{ 用 } \Delta k'=\begin{vmatrix} x_k-a & x_{k+1}-a \\ y_k-b & y_{k+1}-b \end{vmatrix}$$

代替. 而因

$$\Delta k'=\Delta k-\begin{vmatrix} x_k & a \\ y_k & b \end{vmatrix}+\begin{vmatrix} x_{k+1} & a \\ y_{k+1} & b \end{vmatrix},$$

每个相加的行列式在总和中出现两次而且具有相反的符号. 于是我们可以选择原点在多边形的内部. 多边形面积是这些三角形 OP_kP_{k+1} 面积的总和，其中 $k=1,2,\cdots,n(P_{n+1}=P_1)$, 然而 OP_kP_{k+1} 的面积恰好是

$$\frac{1}{2}\begin{vmatrix} x_k & x_{k+1} \\ y_k & y_{k+1} \end{vmatrix}.$$

6. 从头两行中减去第三行，即可见该行列式等于

$$\frac{1}{2}|\mathbf{X}_1\times\mathbf{X}_2|,$$

其中

$$\mathbf{X}_1=(x_1-x_3,y_1-y_3),\ \mathbf{X}_2=(x_2-x_3,y_2-y_3).$$

7. 如果顶点的坐标是有理数，则用行列式定义的三角形的面积显然是有理数. 但是，对于具有边长 s 的等边三角形的面积是 $\frac{1}{4}s^2\sqrt{3}$, 其中

$$s^2=(x_i-x_j)^2+(y_i-y_j)^2 \qquad (i\neq j)$$

显然是有理数.

8. (a) 用向量写法，这说的是

$$\mathbf{A}\cdot(\mathbf{A}'\times\mathbf{A}'')\leq|\mathbf{A}|\cdot|\mathbf{A}'|\cdot|\mathbf{A}''|,$$

这显然是正确的，因为

$$|\mathbf{A}'\times\mathbf{A}''|\leq|\mathbf{A}'|\cdot|\mathbf{A}''|,$$

并且

$$|\mathbf{D}|=|\mathbf{A}\cdot(\mathbf{A}'\times\mathbf{A}'')|\leq|\mathbf{A}|\cdot|\mathbf{A}'\times\mathbf{A}''|.$$

(b) 等号仅当前面两个不等式的等号都成立时才能成立. 因此, $\mathbf{A},\mathbf{A}'$ 和 $\mathbf{A}''$ 必定是互相垂直的.

9. (a) 如果 $\mathbf{B}$ 与 $\mathbf{C}$ 是相关的，譬如说 $\mathbf{C}=\lambda\mathbf{B}$, 这时恒等式显然成立. 否则，设 $\mathbf{B},\mathbf{B}\times\mathbf{C},\mathbf{B}\times(\mathbf{B}\times\mathbf{C})$ 三个方向上的单位向量分别是 $\mathbf{E}_1,\mathbf{E}_2,\mathbf{E}_3$, 则它们形成一组正交基. 通过这组基写出 $\mathbf{A},\mathbf{B},\mathbf{C}$:

$$\mathbf{A}=a_1\mathbf{E}_1+a_2\mathbf{E}_2+a_3\mathbf{E}_3,$$

$$\mathbf{B}=b\mathbf{E}_1,\mathbf{C}=c_1\mathbf{E}_1+c_3\mathbf{E}_3$$

便得到 $\mathbf{B}\times\mathbf{C}=-bc_3\mathbf{E}_2$ 因而

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C})=bc_3(a_3\mathbf{E}_1-a_1\mathbf{E}_3).$$

利用 $\mathbf{E}_1=(1/b)\mathbf{B}$ 与 $\mathbf{E}_3=1/c_3[\mathbf{C}-(c_1/b)\mathbf{B}]$ 便得到

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C})=(a_1c_1+a_3c_3)\mathbf{B}-(a_1b)\mathbf{C}.$$

(b) 注意

$$z=(\mathbf{X}\times\mathbf{Y})\cdot(\mathbf{X}'\times\mathbf{Y}')=\det\,(\mathbf{X},\mathbf{Y},\mathbf{X}'\times\mathbf{Y}')$$

$$=\det(\mathbf{Y},\mathbf{X}'\times\mathbf{Y}',\mathbf{X})=[\mathbf{Y}\times(\mathbf{X}'\times\mathbf{Y}')]\cdot\mathbf{X}.$$

应用练习第 9 题 (a) 便得到

$$z=[(\mathbf{Y}\cdot\mathbf{Y}')\mathbf{X}'-(\mathbf{Y}\cdot\mathbf{X}')\mathbf{Y}']\cdot\mathbf{X}.$$

(c) 应用练习第 9 题 (a) 改写左边表达式为

$$\mathbf{U} = [(\mathbf{X}\cdot\mathbf{Z})\mathbf{Y} - (\mathbf{X}\cdot\mathbf{Y})\mathbf{Z}]\cdot\mathbf{V},$$

其中

$$\begin{aligned}\mathbf{V} &= [(\mathbf{Y}\cdot\mathbf{X})\mathbf{Z} - (\mathbf{Y}\cdot\mathbf{Z})\mathbf{X}] \times [(\mathbf{Z}\cdot\mathbf{Y})\mathbf{X} - (\mathbf{Z}\cdot\mathbf{X})\mathbf{Y}] \\ &= (\mathbf{Y}\cdot\mathbf{X})(\mathbf{Y}\cdot\mathbf{Z})(\mathbf{Z}\times\mathbf{X}) + (\mathbf{X}\cdot\mathbf{Y})(\mathbf{X}\cdot\mathbf{Z})(\mathbf{Y}\times\mathbf{Z}) \\ &\quad + (\mathbf{Z}\cdot\mathbf{Y})(\mathbf{Z}\cdot\mathbf{X})(\mathbf{X}\times\mathbf{Y}).\end{aligned}$$

于是,

$$\begin{aligned}\mathbf{U} =&(\mathbf{X}\cdot\mathbf{Z})(\mathbf{Y}\cdot\mathbf{X})(\mathbf{Y}\cdot\mathbf{Z})[\mathbf{Y}\cdot(\mathbf{Z}\times\mathbf{X})] \\ &- (\mathbf{X}\cdot\mathbf{Y})(\mathbf{Z}\cdot\mathbf{Y})(\mathbf{Z}\cdot\mathbf{X})[\mathbf{Z}\cdot(\mathbf{X}\times\mathbf{Y})] = 0.\end{aligned}$$

10. 设 $\mathbf{E}$ 是方向 $(-1,0,1)$ 上的单位向量, 于是

$$\mathbf{E} = \left(-\frac{\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right).$$

设 $\mathbf{X} = (x, y, z)$ 是任一点的位置向量, 并设从该点到转轴的垂足的位置向量是

$$\mathbf{A} = (\mathbf{X}\cdot\mathbf{E})\mathbf{E} = \left(\frac{1}{2}(x-z), 0, \frac{1}{2}(z-x)\right).$$

注意 $\mathbf{X}-\mathbf{A}$ 是垂直于 $\mathbf{A}$ 的, 并且引进与这两个向量都互相垂直的向量 $\mathbf{E}\times(\mathbf{X}-\mathbf{A})$. 如果 $\mathbf{X}'$ 是 (x, y, z) 旋转所得的像, 则 $\mathbf{X}'-\mathbf{A}$ 也是垂直于 $\mathbf{A}$ 并由所给定的条件给出

$$(\mathbf{X}-\mathbf{A})\times(\mathbf{X}'-\mathbf{A}) = r^2 \sin\phi\mathbf{E},$$

其中 $r = |\mathbf{X}-\mathbf{A}| = |\mathbf{X}'-\mathbf{A}|$ 是 $\mathbf{X}$ 到轴的距离. 令

$$\mathbf{X}' = \lambda\mathbf{A} + \mu(\mathbf{X}-\mathbf{A}) + \nu[\mathbf{E}\times(\mathbf{X}-\mathbf{A})],$$

这是我们可以作的，因为这个线性组合中所出现的向量都是互相垂直的. 从 $(\mathbf{X}'-\mathbf{A})\cdot\mathbf{A}=0$ 得 $\lambda=1$; 从

$$(\mathbf{X}'-\mathbf{A})\cdot(\mathbf{X}-\mathbf{A})=r^2\cos\varphi$$

我们有 $\mu=\cos\varphi$. 最后，根据练习第 9 题 (a),

$$\begin{aligned}r^2\sin\varphi\mathbf{E}&=(\mathbf{X}-\mathbf{A})\times(\mathbf{X}'-\mathbf{A})\\&=V(\mathbf{X}-\mathbf{A})\times[\mathbf{E}\times(\mathbf{X}-\mathbf{A})]\\&=Ur^2\mathbf{E};\end{aligned}$$

所以 $r=\sin\varphi$. 由于

$$\mathbf{X}-\mathbf{A}=\left(\frac{1}{2}(x+z),y,\frac{1}{2}(x+z)\right),$$

$$\mathbf{E}\times(\mathbf{X}-\mathbf{A})=\mathbf{E}\times\mathbf{X}=\frac{1}{2}\sqrt{2}(-y,x+z,-y),$$

便得到 $\mathbf{X}'=\mathbf{a}\mathbf{X}$, 其中

$$\mathbf{a}=\begin{pmatrix}\frac{1}{2}(\cos\phi+1) & -\frac{1}{2}\sqrt{2}\sin\varphi & \frac{1}{2}(\cos\varphi+1)\\ \frac{1}{2}\sqrt{2}\sin\varphi & \cos\varphi & \frac{\sqrt{2}}{2}\sin\varphi\\ \frac{1}{2}(\cos\phi-1) & -\frac{\sqrt{2}}{2}\sin\varphi & \frac{1}{2}(\cos\varphi+1)\end{pmatrix}.$$

11. 根据练习第 9(a) 题,

$$\begin{aligned}\mathbf{X}&=[(\mathbf{A}\times\mathbf{B})\cdot\mathbf{D}]\mathbf{C}-[(\mathbf{A}\times\mathbf{B})\cdot\mathbf{C}]\mathbf{D}\\&=[(\mathbf{C}\times\mathbf{D})\cdot\mathbf{A}]\mathbf{B}-[(\mathbf{C}\times\mathbf{D})\cdot\mathbf{B}]\mathbf{A}.\end{aligned}$$

因为 A,B,C 是无关的， $(\mathbf{A}\times\mathbf{B})\cdot\mathbf{C}\neq 0$, 所以我们可以对 $\mathbf{D}$ 求解.

12. 设 $\mathbf{E}_1', \mathbf{E}_2', \mathbf{E}_3'$ 是新坐标系中的单位坐标向量. 我们有

$$\mathbf{E}_3 \cdot \mathbf{E}_3' = \cos\theta, \mathbf{E}_1 \times (\mathbf{E}_3 \times \mathbf{E}_3') = \sin\theta \sin\varphi \mathbf{E}_3,$$

$$\mathbf{E}_1' \times (\mathbf{E}_3 \times \mathbf{E}_3') = -\sin\theta \sin\psi \mathbf{E}_3'.$$

还有 $\mathbf{E}_1 \cdot (\mathbf{E}_3 \times \mathbf{E}_3') = \sin\theta \cos\phi$ 并且 $\mathbf{E}_1' \cdot (\mathbf{E}_3 \times \mathbf{E}_3') = \sin\theta \cos\psi$. 于是, 从练习第 9(a) 题, 有

$$(\mathbf{E}_1 \cdot \mathbf{E}_3') = \sin\theta \sin\phi, \quad \mathbf{E}_1' \cdot \mathbf{E}_3 = \sin\theta \sin\psi.$$

现在, 令

$$\mathbf{E}_i = \sum_{j=1}^{3} a_{ij} \mathbf{E}_j',$$

其中

$$(a_{ij}) = (\mathbf{E}_i \cdot \mathbf{E}_j')$$

是我们要寻求的方阵. 从我们已经知道的条件便得出

$$a_{13} = \sin\theta \sin\varphi, a_{31} = \sin\theta \sin\psi, a_{33} = \cos\theta,$$

根据 $\mathbf{E}_3 \times \mathbf{E}_3' = \sin\theta \sin\psi \mathbf{E}_2' + a_{32} \mathbf{E}_1'$ 并用 $\mathbf{E}'$ 对其两边取数量积便得到

$$\mathbf{E}_1' \cdot (\mathbf{E}_3 \times \mathbf{E}_3') = \sin\theta \cos\psi = a_{32}.$$

于是

$$\mathbf{E}_3 = -\sin\theta \sin\psi \mathbf{E}_1' + \sin\theta \cos\psi \mathbf{E}_2' + \cos\theta \mathbf{E}_3'.$$

运用这个对于 $\mathbf{E}_3$ 的展开式并在如下方程组中对 a_{11} 和 a_{12} 求解

$$\mathbf{E}_1 \cdot \mathbf{E}_3 = 0, \qquad |\mathbf{E}_1|^2 = 1,$$

便得到

$$a_{11} = -\cos\theta \sin\phi \sin\psi \pm \cos\phi \cos\psi,$$

$$a_{12} = -\cos\theta \sin\phi \cos\psi \pm \cos\phi \sin\psi.$$

在关于 a_{11}, a_{12} 的这些表达式中未确定的符号是通过这一条件

$$\mathbf{E}_1 \cdot (\mathbf{E}_3 \times \mathbf{E}_3') = \sin\theta\cos\phi$$

来确定的，它使得在 a_{11} 表达式中取正号，而在 a_{12} 表达式中取负号．令 $\mathbf{E}_2 = \mathbf{E}_3 \times \mathbf{E}_1$，终于得到

$$(a_{ij}) = \begin{pmatrix} -\cos\theta\sin\phi\sin\psi & -\cos\theta\sin\varphi\cos\psi & \sin\theta\sin\phi \\ +\cos\phi\cos\psi & -\cos\varphi\sin\psi & \\ \cos\theta\cos\phi\cos\psi & \cos\theta\cos\phi\cos\psi & -\sin\theta\cos\phi \\ +\sin\phi\cos\psi & -\sin\phi\sin\psi & \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{pmatrix}.$$

注意这个结果对于 $\theta = 0$ 或 π 也成立，这是在 ϕ 和 ψ 成为不确定的时候，而这分别在 $\phi + \psi = xOx'$ 或 $\phi - \psi = xOx'$ 时发生．

角 ϕ, ψ, θ 也称为欧拉角．我们的结果表明，具有行列式值为 $+1$ 的最一般的正交方阵能够通过三个变量 φ, ψ, θ 表示成“参量形式”，这些参量受下列不等式限制：

$$0 \le \theta \le \pi, \quad 0 \le \phi < 2\pi, \quad 0 \le \psi < 2\pi.$$

13．设 $\mathbf{A} = a_1\mathbf{E}_1 + a_2\mathbf{E}_2 + \cdots + a_m\mathbf{E}_m$ 是 π 上的一个非零向量，它垂直于 π' 上的一切向量，比如说具有 $a_1 \neq 0$．因

$$E_1 = 1/a_1(\mathbf{A} - a_2\mathbf{E}_2 - \cdots - a_m\mathbf{E}_m),$$

根据 (85a) 我们得到

$$\mu = \frac{1}{a_1}[\mathbf{A} - a_2\mathbf{E}_2 - \cdots - a_m\mathbf{E}_m, \mathbf{E}_2, \cdots, \mathbf{E}_m; \mathbf{E}_1', \cdots \mathbf{E}_m']$$

$$= \frac{1}{a_1}[\mathbf{A}, \mathbf{E}_2, \cdots, \mathbf{E}_m; \mathbf{E}_1', \mathbf{E}_2', \cdots, \mathbf{E}_m'] = 0.$$

反之，若 $\mu = 0$，则在 μ 的行列式表示 (85a) 中的列向量是线性相关的；对于某一组不全为零的系数有

$$\lambda_1\mathbf{E}_k \cdot \mathbf{E}_1' + \lambda_2\mathbf{E}_k \cdot \mathbf{E}_2' + \cdots + \lambda_m\mathbf{E}_k \cdot \mathbf{E}_m' = 0$$

$$(k=1,2,\cdots,m),$$

于是

$$\mathbf{E}_k\cdot(\lambda_1\mathbf{E}_1'+\lambda_2\mathbf{E}_2'+\cdots+\lambda_m\mathbf{E}_m')=0,$$

因而我们在 π' 中有一个向量正交于每一个基向量，从而垂直于 π 中的每一个向量.

练习 2.5 (p.231)

1. 设 p 的坐标是 (x_1',x_2',x_3'), Q 的坐标是 (x_1'',x_2'',x_3''). 于是 $\overrightarrow{PQ}$ 代表向量 $\mathbf{U}$, 其中 $u_i=x_i''-x_i'$. P 与 Q 在新坐标系中的坐标由 (89a) 给出 (带有适当的撇 “′”), 因而 $\overrightarrow{PQ}$ 所代表的向量具有分量

$$v_i=y_i''-y_i',$$

它们显然满足 (89a).

5. 设曲线由向量 $\mathbf{X}(t)$ 表示，并设参量的三个值由 t,t_1,t_2 给定. 于是相应的点表示为 $\mathbf{X}=\mathbf{X}(t),\mathbf{X}_1=\mathbf{X}(t_1),\mathbf{X}_2=\mathbf{X}(t_2)$. 通过该三个点的平面的法线平行于

$$(\mathbf{X}_1-\mathbf{X})\times(\mathbf{X}_2-\mathbf{X}).$$

假设 $t_1-t=h_1,t_2-t=h_2$, 并且应用泰勒定理，便得到

$$\mathbf{X}_i=\mathbf{X}+\frac{d\mathbf{X}}{dt}h_i+\frac{1}{2}\frac{d^2\mathbf{X}}{dt^2}h_i^2+\cdots.$$

于是，保留低阶项便得到

$$(\mathbf{X}_1-\mathbf{X})\times(\mathbf{X}_2-\mathbf{X})$$
$$=\frac{1}{2}\frac{d\mathbf{X}}{dt}\times\frac{d^2\mathbf{X}}{dt^2}(hk^2-kh^2).$$

在 h 与 k 趋向于 0 的极限情形下， t 趋向于 t_0, 密切平面在

$$\mathbf{X}_0=\mathbf{X}(t_0)$$

的法线取

$$\frac{d\mathbf{X}}{dt}\times\frac{d^2\mathbf{X}}{dt^2}$$

的方向．于是，密切平面上的每一点的位置向量 $\mathbf{Y}$ 满足

$$(\mathbf{Y}-\mathbf{X}_0)\cdot\left(\frac{d\mathbf{X}}{dt}\times\frac{d^2\mathbf{X}}{dt^2}\right)=0.$$

6. 根据前一练习题的结果，我们必须证明 $\frac{d\mathbf{X}}{ds}$ 与 $\frac{d^2\mathbf{X}}{ds^2}$ 都垂直于 $\frac{d\mathbf{X}}{dt}\times\frac{d^2\mathbf{X}}{dt^2}$. 这是直接根据

$$\frac{d\mathbf{X}}{ds}=\frac{d\mathbf{X}}{dt}\frac{dt}{ds}\ \text{和}\ \frac{d^2\mathbf{X}}{ds^2}=\frac{d\mathbf{X}}{dt}\frac{d^2t}{ds^2}+\frac{d^2\mathbf{X}}{dt^2}\left(\frac{dt}{ds}\right)^2.$$

7. 设曲线由 $\mathbf{X}(s)$ 给定，其中 s 是弧长．用泰勒定理展开 $\mathbf{X}$:

$$\mathbf{X}(s)=\mathbf{X}(s_0)+\mathbf{X}'(s_0)l+\mathbf{Y}O(l^2),$$

其中 $l=s-s_0$ 而 $\mathbf{Y}$ 是有界的．于是，因为 $|\mathbf{X}'(s_0)|=1$, 我们有

$$\begin{aligned}d-l&=|\mathbf{X}(s)-\mathbf{X}(s_0)|-l\\&=|\mathbf{X}'(s_0)l-\mathbf{Y}O(l^2)|-l\\&\le|\mathbf{X}'(s_0)|l+O(l^2)-l;\end{aligned}$$

这意味着 $d-l=O(l^2)=o(l)$.

8. 根据上面第 6 题的解答，有

$$k=\left|\frac{d^2\mathbf{X}}{ds^2}\right|=\left|\mathbf{X}'\frac{d^2t}{ds^2}+\mathbf{X}''\left(\frac{dt}{ds}\right)^2\right|.$$

注意到

$$\frac{dt}{ds}=\frac{1}{|\mathbf{X}'|},$$

即有

$$\frac{d^2t}{ds^2}=-\frac{\mathbf{X}'\cdot\mathbf{X}''}{|\mathbf{X}'|^4};$$

于是

$$k^2 = \frac{|\mathbf{X}'|^2|\mathbf{X}''|^2 - (\mathbf{X}' \cdot \mathbf{X}'')^2}{|\mathbf{X}'|^6}.$$

9. 根据练习第 6 题的解答， $\dfrac{d^2\mathbf{X}}{dt^2}$ 是 $\dfrac{d\mathbf{X}}{ds}$ 与 $\dfrac{d^2\mathbf{X}}{ds^2}$ 的线性组合.

10. 设 C 用 $\mathbf{X}(t)$ 表示，并假设 $\mathbf{B}$ 的位置向量 $\mathbf{X}(t_0)$ 不是 C 的端点. 设 $\mathbf{Y}$ 是 A 的位置向量. $|\mathbf{Y} - \mathbf{X}(t_0)|$ 是极小，条件是

$$\frac{d}{dt}|\mathbf{Y} - \mathbf{X}(t)|^2_{t=t_0} = 0;$$

这也就是

$$[\mathbf{Y} - \mathbf{X}(t_0)] \cdot \mathbf{X}'(t_0) = 0.$$

11. 设曲线由含参量的向量 $\mathbf{X}(\theta)$ 给定，其中 $x = a\cos\theta$, $y = a\sin\theta$. 切平面只与 x 和 y 有关而与 z 无关，并且它与 y 轴的夹角为 θ. 曲线的切向量 $\mathbf{X}'$ 的 z 分量满足

$$\frac{z'}{\sqrt{x'^2 + y'^2 + z'^2}} = \cos\theta,$$

或

$$\frac{z'}{\sqrt{a^2 + z'^2}} = \cot\theta.$$

于是

$$z' = \pm a\cot\theta;$$

从而

$$z = c \pm a\log\sin\theta.$$

关于该曲线的曲率参看练习第 8 题.

12. 根据 $d\mathbf{X}/d\theta = (-\sin\theta, \cos\theta, \sinh A\theta)$, 我们有

$$\frac{d^2\mathbf{X}}{d\theta^2} = (-\cos\theta, -\sin\theta, A\cosh A\theta),$$

解答给出对于密切平面上任何一点 $\mathbf{Y}$ 的方程

$$0 = (\mathbf{Y} - \mathbf{X}) \cdot \left(\frac{d\mathbf{X}}{d\theta} \times \frac{d^2\mathbf{X}}{d\theta^2}\right),$$

其中法向量给定如下:

$$\frac{d\mathbf{X}}{d\theta} \times \frac{d^2\mathbf{X}}{d\theta^2} = (N_1, N_2, N_3),$$

$$N_1 = A\cos\theta\cosh A\theta + \sin\theta\sinh A\theta,$$

$$N_2 = A\sin\theta\cosh A\theta - \cos\theta\sinh A\theta,$$

$$N_3 = 1.$$

原点到平面的距离是 $|\mathbf{X}\cdot\mathbf{N}|/|\mathbf{N}|$, 而由于 $\mathbf{X}\cdot\mathbf{N} = (A+1/A)\cosh A\theta$ 且 $|\mathbf{N}|^2 = (A^2+1)\cosh^2 A\theta$, 便立即得到结果.

13. (a). 设 $\mathbf{X}(t)$ 是曲线的参量表达式, 并令 $\mathbf{X}_i = \mathbf{X}(t_i)$. 根据 2.4 节练习第 3 题, 通过这三个点的平面是

$$(\mathbf{X}_1 - \mathbf{X}) \cdot [(\mathbf{X}_2 - \mathbf{X}) \times (\mathbf{X}_3 - \mathbf{X})] = 0,$$

或

$$\mathbf{X} \cdot [\mathbf{X}_1 \times \mathbf{X}_2 + \mathbf{X}_2 \times \mathbf{X}_3 + \mathbf{X}_3 \times \mathbf{X}_1] = \mathbf{X}_1 \cdot (\mathbf{X}_2 \times \mathbf{X}_3),$$

因此结果成立.

(b) 该三个密切平面有方程

$$(\mathbf{X} - \mathbf{X}_i) \cdot (\mathbf{X}_i' \times \mathbf{X}_i'') = 0$$

(根据练习第 6 题), 或借助坐标写出如下:

$$\frac{3x}{a} - \frac{6t_i}{b}y + \frac{3t_i^2}{c}z - t_i^3 = 0.$$

于是, 如果 (x, y, z) 是三个密切平面的一个公共点, 则 t_1, t_2, t_3 是上述方程的三个根, 该方程的系数是

$$t_1 + t_2 + t_3 = \frac{3z}{c},$$

$$t_1t_2 + t_2t_3 + t_3t_1 = \frac{6y}{b},$$

$$t_1t_2t_3 = \frac{3x}{a}.$$

14. 因为一个球面决定于不共面的任意四个点，我们可以附加四个条件在密切球上；球面与曲线的密切接触是 3 阶的．设 $\mathbf{X}(s)$ 是曲线用弧长作参量的表达式，而 A 是球的重心．要求 $(\mathbf{X}-\mathbf{A})$ 是 3 阶地趋于 0; 于是，从 $|\dot{\mathbf{X}}|^2 = 1$ 与 $\dot{\mathbf{X}}\cdot\ddot{\mathbf{X}} = 0$,

$$(\mathbf{X}-\mathbf{A})\cdot\dot{\mathbf{X}} = 0,$$

$$(\mathbf{X}-\mathbf{A})\cdot\ddot{\mathbf{X}} + 1 = 0,$$

$$(\mathbf{X}-\mathbf{A})\cdot\dddot{\mathbf{X}} = 0.$$

从这三个方程的头一个和最后一个得到

$$\mathbf{X}-\mathbf{A} = \lambda(\dot{\mathbf{X}}\times\dddot{\mathbf{X}}),$$

其中 λ 是通过第二个方程给定的．因此

$$\mathbf{A} = \mathbf{X} + \frac{\dot{\mathbf{X}}\times\dddot{\mathbf{X}}}{\ddot{\mathbf{X}}\cdot[\dot{\mathbf{X}}\times\dddot{\mathbf{X}}]}.$$

15. 在上题的解答中令 $|\mathbf{X}-\mathbf{A}| = 1$.

16. 根据练习第 6 题，因为 ξ_3 是切平的法线，所以 $\frac{1}{\tau} = |\xi_3|$. 其次，因为 $\dot{\xi}_i$ 与 ξ_i 是互相垂直的，

$$\dot{\xi}_2 = a\xi_1 + b\xi_3,$$

$$\dot{\xi}_3 = c\xi_1 + d\xi_2.$$

微商 $\xi_1 = \xi_2\times\xi_3$, 得到

$$\frac{1}{\rho}\xi_2 = (\xi_2\times\dot{\xi}_3) + (\dot{\xi}_2\times\xi_3) = -a\dot{\xi}_2 - c\xi_3;$$

因此 $a=-1/\rho$, $C=0$. 根据 $\dot{\xi}_3=d\xi_2$ 有 $d=\pm 1/\tau$; 取负号. 为了确定 b, 微商 $\xi_3=(\xi_1\times\xi_2)$:

$$\dot{\xi}_3=-\frac{1}{\tau}\xi_2=(\xi_1\times\dot{\xi}_2)-(\xi_2\times\dot{\xi}_1)=-b\xi_2;$$

因此 $b=1/\tau$.

17. (a) 微分 $\ddot{\mathbf{X}}=\dot{\xi}_1=k\xi_2$, 得到

(a) $\dddot{\mathbf{X}}=K\xi_2+K\dot{\xi}_2=-k^2\xi_1+\dot{k}\xi_2+\dfrac{k}{2}\xi_3$.

(b) 根据练习第 14 题的结果, 得到

$$\frac{\xi_2}{\tau}+\frac{\dot{k}}{k^2\tau}\xi_3.$$

18. 因为 $1/\tau=|\dot{\xi}_3|=0$, 所以 $\dot{\xi}_3=0$ 并因此 ξ_3 必然是一个常向量. 于是由 $0=\xi_1\cdot\xi_3=\dot{\mathbf{X}}\cdot\xi_3=\dfrac{d}{ds}(\mathbf{X}\cdot\xi_3)$ 可以推知 $\mathbf{X}\cdot\xi_3=$ 常数.

19. 设 $\mathbf{A}$ 与 $\mathbf{P}$ 分别是 A 与 P 的位置向量. 令 $\mathbf{X}=\mathbf{A}-\mathbf{P}$, 因而 $\dot{\mathbf{X}}=-\dot{\mathbf{P}}$. 这个等式给出

$$\frac{d}{dt}|\mathbf{X}|=-\mathbf{a}\cdot\dot{\mathbf{P}},$$

因为直接从微分公式得出

$$\frac{d}{dt}|\mathbf{X}|=\frac{d}{dt}\sqrt{\mathbf{X}\cdot\mathbf{X}}=\frac{\mathbf{X}\cdot\dot{\mathbf{X}}}{|\mathbf{X}|},$$

只需令 $\mathbf{a}=\mathbf{X}/|\mathbf{X}|$.

20. (a) 像上题的解答中那样令 $\mathbf{X}=\mathbf{A}-\mathbf{P}$. 根据上题的解答有

$$-\dot{\mathbf{P}}=\dot{\mathbf{X}}=\frac{d}{dt}(|\mathbf{X}|\mathbf{a})=-(\mathbf{a}\cdot\dot{\mathbf{P}})\mathbf{a}+|\mathbf{X}|\dot{\mathbf{a}};$$

从而直接得出所要的结果.

(b) 在下式中引进 $\dot{\mathbf{a}}$ 的表达式以及 $\dot{\mathbf{b}}$ 的类似的表达式:

$$\ddot{\mathbf{P}}=u\dot{\mathbf{a}}+v\dot{\mathbf{b}}+w\dot{\mathbf{c}}+\dot{u}\mathbf{a}+\dot{v}\mathbf{b}+\dot{w}\mathbf{c}.$$

21. (a) 设曲线通过 $\mathbf{X}(t)$ 给定. 则曲面有参量方程

$$\mathbf{Y}=\mathbf{X}(t)+\lambda\dot{\mathbf{X}}(t).$$

向量 $\dfrac{\partial\mathbf{Y}}{\partial\lambda}\times\dfrac{\partial\mathbf{Y}}{\partial t}$ 是曲面的法线, 但是

$$\frac{\partial\mathbf{Y}}{\partial\lambda}\times\frac{\partial\mathbf{Y}}{\partial t}=\dot{\mathbf{X}}(t)\times[\dot{\mathbf{X}}(t)+\lambda\ddot{\mathbf{X}}(t)]=\lambda\dot{\mathbf{X}}(t)\times\ddot{\mathbf{X}}(t)$$

也是密切平面的法线.

(b) 令 $\mathbf{Y}=(x,y,z)$, $\mathbf{X}(t)=(\alpha(t),\beta(t),\gamma(t))$, 于是, x 与 y 是 t 与 λ 的函数并且满足

$$x=\alpha(t)+\lambda\dot{\alpha}(t),$$

$$y=\beta(t)+\lambda\dot{\beta}(t).$$

用

$$u(x,y)=\gamma(t)+\lambda\dot{\gamma}(t)$$

通过对 t 与 λ 的微商来计算 u_{xx},u_{yy},u_{xy}. 对 x 微商

$$\mathbf{Y}=\mathbf{X}(t)+\lambda\dot{\mathbf{X}}(t)$$

得到 $(\lambda=s)$

$$\mathbf{Y}_x=(1,0,u_x)=(\dot{\mathbf{X}}+\lambda\ddot{\mathbf{X}})t_x+\dot{\mathbf{X}}_{sx}.$$

写出 $\dot{\mathbf{X}}\times\mathbf{Y}_x$ 并在 x 与 z 方向上让分量相等便得到

$$\dot{\beta}u_x=st_x(\beta,\gamma),\qquad\dot{\beta}=-st_x(\alpha,\beta),$$

其中 (u,v) 是如下定义的

$$(u,v)=\dot{u}\ddot{v}-\dot{v}\ddot{u}.$$

于是

$$u_x=-\frac{(\beta,\gamma)}{(\alpha,\beta)},\quad t_x=-\frac{\dot{\beta}}{s(\alpha,\beta)}.$$

类似地，从 $\dot{\mathbf{X}} \times \mathbf{Y}_y$ 得到

$$u_y = -\frac{(\gamma,\alpha)}{(\alpha,\beta)}, \quad t_y = \frac{\dot{\alpha}}{s(\alpha,\beta)}.$$

注意 u_x 与 u_y 不依赖于 λ, 便由此得出

$$u_{xx} = t_x \frac{d}{dt} u_x = \frac{\dot{\beta}}{s(\alpha,\beta)} \frac{d}{dt} \frac{(\beta,\gamma)}{(\alpha,\beta)},$$

$$u_{yy} = t_y \frac{d}{dt} u_y = \frac{\dot{\alpha}}{s(\alpha,\beta)} \frac{d}{dt} \frac{(\alpha,\gamma)}{(\alpha,\beta)},$$

$$u_{xy} = t_y \frac{d}{dt} u_x = -\frac{\dot{\alpha}}{s(\alpha,\beta)} \frac{d}{dt} \frac{(\beta,\gamma)}{(\alpha,\beta)}$$

$$= t_x \frac{d}{dt} u_y = -\frac{\dot{\beta}}{s(\alpha,\beta)} \frac{d}{dt} \frac{(\alpha,\gamma)}{(\alpha,\beta)}.$$

由此直接得出结果.

第三章

练习 3.1 a (p.237)

1. 令 $y_{n+1} = Y_n + Cf(a, y_n)$, 其中 C 是常数. 然后用第一卷第 6.3 节的 c 与 d, 取 $\varphi(y) = Y + Cf(a, y)$. 为了保证收敛，我们要求 $|\varphi'(y)| \leq q < 1$ 在某个包含 b 的区间上，并取足够小的 q. 因此，我们试图固定 C 使得 $\varphi'(y)$ 接近于 0, 或

$$C \approx -\frac{1}{f_y(a,b)}.$$

因此我们先假设 $f_y(a,b) \neq 0$.

实际上，我们取 $C = -1/f_y(a, y_0)$, 其中 y_0 接近于要寻找的解 b. 于是收敛的条件成为

$$|\varphi'(y)| = \left|\frac{f_y(a,y_0) - f_y(a,y)}{f_y(a,y_0)}\right| \leq q < 1,$$

对 b 的某个邻域内的一切 y 成立. 假设 f_y 满足利普希茨条件, 即不等式

$$|f_y(a,\eta_2) - f_y(a,\eta_1)| < k|\eta_2 - \eta_1|$$

在 b 的某个邻域成立. 在这个邻域内, 设 ϵ 是某个也许小一些的邻域的半径, 在其中 $\partial f/\partial y$ 离开 0 有一个距离, 即

$$f_y(a,y) > m > 0;$$

这样的邻域确是存在的, 这是由于利普希茨条件以及

$$f_y(a,b) \neq 0,$$

无论初始选择怎样的 y_0 只要满足

$$|y_0 - b| < \max\left\{\epsilon, \frac{qm}{2K}\right\},$$

就有迭代过程收敛到 b, 这是因为

$$|Y_n - b| \leq \frac{1}{2}q^n|Y_0 - b|.$$

练习 3.1 b (p.239)

1. (a) 切平面是水平的. 曲面交切平面于一对直线 $y = x$ 与 $y = -x$; 因此, 在 (x_0, y_0) 邻域内, y 不可能表示成 x 的函数.

(b) 曲面是母线平行于向量 $\mathbf{i} - \mathbf{j}$ 的柱面, 于是直线

$$y = 1 - x, \qquad z = 0$$

落在曲面上并且给出所需要的解 $y = 1 - x$.

(c) 曲面是母线平行于 $\mathbf{i} - \mathbf{j}$ 的柱面, 解是 $y = 1/2 - x$.

(d) 切平面 $y + z = 0$ 不是水平的. 因此曲线 $f(x,y) = 0$ 在原点与直线 $y = 0$ 相切.

练习 3.1 c (p.243)

1. 从两边减去右边的常数，便可将每个方程表示成

$$F(x,y)=0$$

的形式. 定理的条件是满足的. 特别是，每个满足 $F(x_0,y_0)=0$ 的给定点是初始解. 而且 $F_y(x_0,y_0)$ 有非零值，即 (a) 4, (b) -1, (c) 2, (d) 6.

2. (a) $-\dfrac{2x+y}{x+2y}$; $-\dfrac{5}{4}$.

(b) 显式，$y=\pi/2x$; 因此 $y'=-\pi/2x^3$.

隐式 $y'=\dfrac{\cot xy-xy}{x^2}$; $-\dfrac{\pi}{2}$.

(c) 显式 $y=1/x$; 因此 $y'=-1/x^2$.

隐式 $y'=-y/x$; -1.

(d) $y'=-\dfrac{y+5x^4}{x+5y^4}$; -1.

3. (a) $y''=\dfrac{-6(x^2+xy+y^2)}{(x+2y)^3}=\dfrac{-42}{(x+2y)^3}$; $-\dfrac{21}{32}$.

(b) $y''=\dfrac{\pi}{x^3}$; π.

(c) $y''=\dfrac{2y}{x^2}=\dfrac{2}{x^3}$; 2.

(d) $y''=\dfrac{-[150x^3y^3(10-xy)+20(x^6+y^6)+8xy-30]}{(x+5y^4)^3}$;

$-\dfrac{19}{3}$.

4. 根据它的二阶微商的正号，(b) 与 (c).

5. 假设在每一个取极值的邻域里，所给方程确定 y 作为 x 的一个可微函数，则在极值点有 $F_x(x,y)=0$. 极大值 $y=b$; 极小值 $y=-b$.

6. 令 $F(x,y)=y-y_0-\displaystyle\int_{x_0}^{x}f_y(\xi,y)d\xi$ 并注意

$$F_y(x,y)=1-\int_{x_0}^{x}f_y(\xi,y)d\xi>0$$

对于足够接近 x_0 的 x 成立.

练习 3.1 d (p.247)

1. 在 (0,0) 附近 $f(x,y)=y^3+x$.

2. 就给出与练习第 1 题相同的函数.

3. 因为 $F_y(x,y)=(3y^2-2y+1)+x^2$ 是关于 y 的二次恒正表达式与一个平方的和, 故有 $F_y(x,y)>0$, 对一切 x,y 成立. 所以对每一个 $x, F(x,y)$ 是对 y 严格增加的, 从而 $F(x,y)=0$ 对于每一个固定的 x, 不可能有多于一个的解. 这样的解必定存在, 因为, 对于每一个 x, $y^3-y^2+(1+x^2)y=G(x,y)$ 可取正、负两种符号的任意大的值, 只要 y 取适当的值. 根据中值定理, 可见 $G(x,y)$ 取遍一切实数值. 特别是, 可找到 y 的某个值使得 $G(x,y)=\phi(x)$; 因此, 对每一个 x 和这个 y 值, 我们有

$$F(x,y)=G(x,y)-\phi(x)=0.$$

练习 3.1 e (p.249)

1. 令 $F(x,y,z)=x+y+z-\sin xyz$, 经计算知

$$F_z(0,0,0)=1\neq 0,$$

$$\frac{\partial z}{\partial x}=\frac{yz\cos xyz-1}{1-xy\cos xyz},\frac{\partial z}{\partial y}=\frac{xz\cos xyz-1}{1-xy\cos xyz}.$$

2. 因为每一个方程都能表达成形如 $F(z,x,y,\cdots)=0$, 其中 F 是经过有理运算并应用单变量连续可微函数形成的, 所以只需检验在该点处的微商 F_z 不为 0.

(a) $F_z=1$,

(b) $F_z=-6$,

(c) 对于

$$F(x,y,z)=1+x+y-\cosh(x+z)-\sinh(y+z),$$

$$F_z = 1.$$

3. 对于

$$f(x,y,z) = x + y + z + xyz^3, \qquad f_z(0,0,0) = 1 \neq 0,$$

第二阶到第四阶项全等于 0; $z = -x - y + \cdots$.

练习 3.2 a (p.255)

1. (a) 方程只在点 (0,0) 被满足; 切线和法线都不存在.

(b) $(\xi - x)[e^x \sin y - e^y \sin x] + (\eta - y)[e^x \cos y + e^y \cos x] = 0$;

$(\eta - x)[e^x \cos y + e^y \cos x] - (\eta - y)[e^x \sin y - e^y \sin x] = 0.$

(c) 方程只在点 $(-1, \pi/2 + 2k\pi)$ 被满足; 切线和法线都不存在.

(d) $(\xi - x)(2x + \cos x) + (\eta - y)(2y - 1) = 0$;

$(\xi - x)(2y - 1) - (\eta - y)(2x + \cos x) = 0.$

(e) $(\xi - x)(3x^2) + (\eta - y)(4y^3 - \sinh y) = 0$;

$(\xi - x)(4y^3 - \sinh y) - (\eta - y)(3x^2) = 0.$

(f) 方程只在正 x 轴和正 y 轴被满足. 对于 $x = 0$, $y > 0$ 切线是 $x = 0$ 而法线是 $\eta = y$; 对于 $y = 0$, $x > 0$, 切线是 $y = 0$ 而法线是 $\xi = x$.

2. -1.

3. 根据第一卷 (p.488) 4.1h 的问题 5,

$$\kappa = \frac{r^2 + 2r'^2 - rr''}{(r^2 + r'^2)^{3/2}},$$

其中撇 "′" 标志着对 θ 的微商. 在 κ 的公式中, 通过 f 的偏微商. 写入 r' 和 r'' 的表达式便得到

$$\kappa = \frac{r^2 f_r^3 + r(f_r^2 f_{\theta\theta} - 2 f_\theta f_r f_{r\theta} + f_\theta^2 f_{rr}) + 2 f_\theta^2 f_r}{(f_\theta^2 + r^2 f^2)^{3/2}}$$

4. 注意 $F_{xx}=F_{yy}=6(x+y-a)=0$ 当 $x+y=a$ 应用 (13), 便得到

$$F_y^2F_{xx}-2F_xF_{xy}+F_x^2F_{yy}=-54xyF_{xy}=0,$$

因为在交线上有 $xy=0$.

5. $a=\pm1$, $b=-\dfrac{1}{2}$.

6. 圆周 K, K', K'' 可以用下列方程表示:

$$K=x^2+y^2+ax+by+c=0,$$

$$K'=x^2+y^2+a'x+b'y+c'=0,$$

$$K''=x^2+y^2+a''x+b''y+c''=0.$$

因此, 经过 A 和 B 的任一个圆周通过 $K'+\lambda K''=0$ 给定. 要圆周 K 正交于 K' 与 K'' 的条件是

$$aa'+bb'-2(c+c')=0,$$

$$aa''+bb''-2(c+c'')=0.$$

根据这些条件, 立即可得表达 K 与 $K'+\lambda K''$ 正交性的相应关系.

练习 3.2 b (p.258)

1. (a). 二重点;

(b) 两个分支切于 x 轴;

(c). 一个角点 —— 对于 $x=0^+$ 斜率是 0, 对于 $x=0^-$ 斜率是 1;

(d) 尖点;

(e) 尖点.

2. 坐标轴.

3. $y=x^2(1\pm x^{1/2})$. 曲线在原点形成一个尖点, 它的两个分支位于它们公切线的同一侧.

4. 曲线族是由曲线 $(x-b)^3 = cy^2$ 经过角度 α 的旋转得到的.

5. 对方程 $F=0$ 关于 x 微商两次并应用 $F_y=0$ 这个事实，便得到

$$\varphi = \arctan \frac{2\sqrt{F_{xy}^2 - F_{xx}F_{yy}}}{F_{xx}+F_{yy}};$$

于是有

(a) $\pi/2$; (b) $\pi/2$.

6. 注意在原点的切线是 $y=0$ 与 $ax+by=0$, 分别在两种情况下展开 y 到二阶:

$$y = \frac{1}{2}y_0''x^2 + \cdots,$$

$$y = -\frac{a}{b}x + \frac{1}{2}y_0''x^2 + \cdots.$$

将这些表达式写入原始方程里，便得到 y_0''. 于是得到

$$k = \frac{2c}{a},\ k = \frac{2(a^3g - a^2bf - ab^2e - b^3c)}{a(a^2+b^2)^{3/2}}.$$

练习 3.2 c (p.261)

1. (a) $5x + 7y - 21z + 9 = 0$,

   (b) $20x + 13y + 3z = 36$,

   (c) $x - y - z + \pi/6 = 0$,

   (d) $x + 2z - 2 = 0$,

   (e) 该曲面在原点没有切平面,

   (f) $z = 0$.

2. 每个方程都有这样一个形式: $F(x,y,z) =$ 常数. 垂直于各个曲面的向量 (F_x, F_y, F_z) 给出如下:

$$\left(\frac{y}{z}, \frac{x}{z}, -\frac{xy}{z^2}\right),$$

$$\left(\frac{x}{\sqrt{x^2+z^2}}, \frac{y}{\sqrt{y^2+z^2}}, \frac{z}{\sqrt{x^2+z^2}}+\frac{z}{\sqrt{y^2+z^2}}\right),$$

$$\left(\frac{x}{\sqrt{x^2+z^2}}, -\frac{y}{\sqrt{y^2+z^2}}, \frac{z}{\sqrt{x^2+z^2}}-\frac{z}{\sqrt{y^2+z^2}}\right).$$

这些向量的任何两个的数量积都等于 0.

3. $x(y+z)=ay$.

4. 因为这是一个旋转曲面，我们可以假定 $y=0$. 设 $(a,0,c)$ 是这曲面上的点，因而 $a^2-c^2=1$. 在这点的切平面是

$$ax-cz=1.$$

交线是 $(z-c)c=(x-a)a=\pm acy$.

5. 根据欧拉关系，切平面的方程

$$(\xi-x)F_x+(\eta-y)F_y+(\zeta-z)F_z=0$$

能表成如下形式:

$$\xi F_x+\eta F_y+\zeta F_z=xF_x+yF_y+zF_z=hF(x,y,z)=h.$$

6. $z_x=\dfrac{yz-x^2}{z^2-xy}, z_y=\dfrac{xz-y^2}{z^2-xy}$.

7. (a) 0,

(b) arc cos $1/\sqrt{6}$,

(c) arc cos $4/5$,

(d) $\pi/2$,

(e) 没有意义.

练习 3.3 a (p.268)

1. (a) 圆 $\xi^2+\eta^2=e^{2x}$; 通过原点的直线 $\xi\sin y-\eta\cos y=0$;

(b) 抛物线弧, $\eta=\sqrt{x^2-2\xi x}$, $\eta=\sqrt{y^2+2\xi y}$;

(c) $\eta=\cos x(1+1/\xi^2)$, $\eta=\cos y(1+\xi^2)$;

(d) 抛物线

$$\xi = \eta^2 - 2\eta(1+x^2) + x^4 + 3x + 1,$$

$$\eta = \xi^2 - 2\xi y + y^4 + y + 1;$$

(e) $\xi = x^{\eta^{1/x}}, \eta = y^{\xi^{1/y}}$;

(f) 直线 $\xi =$ 常数, $\eta =$ 常数 $(\eta \geq 1)$;

(g) 椭圆弧

$$\xi^2 - 2\xi\eta \sin 2x + \eta^2 = \cos^2 2x,$$

$$\xi^2 - 2\xi\eta \sin 2x + \eta^2 = \cos^2 2y;$$

(h) 线段 $\xi = e^{\cos x}$, $(e^{-1} \leq \eta \leq e)$, $\eta = e^{\cos y}$, $(e^{-1} \leq \xi \leq e)$.

2. 这方程只允许在 $x = y = 0$ 取值. 所以区域是平面, 它的象是 $\xi\eta$ 平面中的开的第一像限.

3. 区域是由两个圆 $\xi^2 + \eta^2 = 8, \xi^2 + \eta^2 = 32$ 以及两条双曲线 $\xi^2 - \eta^2 = 2, \xi^2 - \eta^2 = 6$ 围成的.

4. 不是. 这里 $\xi\eta$ 平面的原点是任何一个点 $(0, y)$ 的像.

练习 3.3 b (p.271)

1. 为此, 只需证明, 在具有笛卡儿坐标 (a, b) 的点上, 曲线 $\xi = \alpha, \eta = \beta$ 有不同的方向, 其中

$$\alpha = (\sin b)/(a-1), \quad \beta = a \tan b.$$

对于 $\xi = \alpha$ 有

$$\frac{dx}{dy} = \frac{(a-1)\cos b}{\sin b};$$

对于 $\eta = \beta$ 有

$$\frac{dx}{dy} = \frac{-a}{\cos^2 b \sin b}.$$

于是, 除了满足 $\cos^3 b = a/(1-a)$ 的点外, 曲线坐标对一切点都有定义.

2. $(\xi-1)^{2/3}+\eta^2(\xi-1)^{-2/3}=1$.

3. 如在第 1 题的解答中那样，这些点具有笛卡儿坐标 (a,b)，在这些点曲线 $\xi=\alpha$ 和 $\eta=\beta$ 有同一个方向，在这种情况下，这些点在 "45° 直线"$b=\pm a$ 上.

练习 3.3 c (p.274)

1. 用

$$\xi^2+\eta^2+\zeta^2=(x^2+y^2+z^2)^{-1}$$

就得到

$$x=\frac{\xi}{\xi^2+\eta^2+\zeta^2},y=\frac{\eta}{\xi^2+\eta^2+\zeta^2},z=\frac{\zeta}{\xi^2+\eta^2+\zeta^2}.$$

2. $r=\sqrt{x^2+y^2+z^2+w^2}$,

$$\phi=\arctan\frac{\sqrt{x^2+y^2+z^2}}{w},\quad \psi=\arctan\frac{\sqrt{y^2+z^2}}{x},$$

$$\theta=\arctan z/y.$$

这里 $r=$ 常数是三维球面，其中心在原点半径为 r; $\phi=$ 常数是超锥面，由经过 O 并与 w 轴成 ϕ 角的一切直线产生的; $\psi=$ 常数的集合是通过 w 轴并与 x 轴交成 ψ 角的所有平面的和集; $\theta=$ 常数的集合是包含 x 轴， w 轴并与 y 轴夹 θ 角的所有三维空间的和集.

练习 3.3 d (p.278)

1. (a) $ad-bc$, (b) $1/\sqrt{x^2+y^2}$ (c) $4xy$
   (d) $1/(x^2+y^2)$ (e) $-3x^2y^2$ (f) $9x^2y^2+1$.

2. 若 $ad-bc=0$, 所有点; 若 $ad-bc\neq0$, 没有.

(b) 没有 (对于 $x=y=0$ 变换没有意义).

(c) 坐标轴.

(d) 没有、可是，因为所有点 $(x,y+2n\pi)$ 有同一的像所以没有整体的逆.

(e) 坐标轴.

(f) 没有.

3. (a) $D = e^{2x}; x_\xi = y_\eta = \xi/(\xi^2+\eta^2)$;

$x_\eta = -y_\xi = \eta/(\xi^2+\eta^2)$;

$x_{\xi\xi} = y_{\xi\eta} = -x_{\eta\eta} = (\xi^2-\eta^2)/(\xi^2+\eta^2)^2$;

$y_{\xi\xi} = -x_{\xi\eta} = -y_{\eta\eta} = -2\xi\eta/(\xi^2+\eta^2)^2$.

(b) $D = 4(x^2+y^2)$; 满足 $r = \sqrt{\xi^2+\eta^2}$, $0 = \arctan \eta/\xi$;

$x_\xi = y_\eta = \frac{1}{2}\sqrt{r}\cos\frac{1}{2}\theta$;

$y_\xi = -x_\eta = -\frac{1}{2}\sqrt{r}\sin\frac{1}{2}\theta$;

$x_{\xi\xi} = y_{\xi\eta} = -x_{\eta\eta} = -\frac{1}{4}r^{3/2}\cos 3\theta/2$;

$y_{\xi\xi} = -x_{\xi\eta} = -y_{\eta\eta} = \frac{1}{4}r^{3/2}\sin 3\theta/2$.

(c) $D = 2\sin(x-y)/\cos^2(x+y)$, $x_\xi = y_\xi = 1/2(1+\xi^2)$;

$x_\eta = y_\eta = 1/2\sqrt{1-\eta^2}$; $x_{\xi\xi} = y_{\xi\xi} = -\xi/(1+\xi^2)^2$;

$x_{\xi\eta} = y_{\xi\eta} = 0$; $x_{\eta\eta} = -y_{\eta\eta} = \eta/2(1-\eta^2)^{3/2}$.

(d) $D = \cosh(x+y)$; $x_\xi = (\cosh y)/D$; $x_\eta = -(\sinh y)/D$;

$y_\xi = (\sinh x)/D$; $y_\eta = (\cosh x)/D$.

$x_{\xi\xi} = -[\cosh^2 y\sinh(x+y) + \sinh^2 x]/D^3$;

$x_{\xi\eta} = \frac{1}{2}[\sinh 2y\sinh(x+y) - \sinh 2x]/D^3$;

$x_{\eta\eta} = -[\sinh^2 y\sinh(x+y) + \cosh^2 x]/D^3$;

$y_{\xi\xi} = [\cosh^2 y - \sinh^2 x\sinh(x+y)]/D^3$;

$y_{\xi\eta} = -\frac{1}{2}[\sinh 2y + \sinh 2x\sinh(x+y)]/D^3$;

$y_{\eta\eta} = [\sinh^2 y - \cosh^2 x\sinh(x+y)]/D^3$.

(e) $D = 6x^3y - 3y^4$, $x_\xi = 2x/3(2x^3-y^3)$,

$x_\eta = -y/(2x^3-y^3)$, $y_\xi = -y/3(2x^3-y^3)$,

$y_\eta = x^2/y(2x^3-y^3)$;

$x_{\xi\xi} = -\frac{2}{3}x(8x^3+5y^3)/(2x^3-y^3)^3$,

$x_{\xi\eta} = 2y(7x^3+y^3)/3(2x^3-y^3)^3$,

$x_{\eta\eta} = -2x^2(x^3+4y^3)/y(2x^3-y^3)^3$,

$$y_{\xi\xi} = 2y(7x^3 + y^3)/3(2x^3 - y^3)^3,$$
$$y_{\xi\eta} = -2x^2(x^3 + 4y^3)/3y(2x^3 - y^3)^3,$$
$$y_{\eta\eta} = 2x(y^6 + 3x^3y^3 - x^6)/y^3(2x^3 - y^3)^3.$$

4. (a) 设 m_1 和 m_2 是 x, y 平面上通过点 (a, b) 的两条曲线的斜率. 设 μ_1 和 μ_2 是 ξ, η 平面上对应点的对应斜率. 用

$$\mu = \frac{d\eta}{d\xi} = \frac{d\eta/dx}{d\xi/dx} = \frac{(\partial\eta/\partial x) + m(\partial\eta/\partial y)}{(\partial\xi/\partial x) + m(\partial\xi/\partial y)}$$
$$= \frac{m(a^2 - b^2) - 2ab}{b^2 - a^2 - 2mab}$$

得

$$\frac{\mu_2 - \mu_1}{1 + \mu_1\mu_2} = \frac{m_1 - m_2}{1 + m_1m_2}.$$

因此, 两个曲线间的夹角数量上保持一致而转向相反.

(b) 注意 $\xi^2 + \eta^2 = 1/(x^2 + y^2)$. 将圆

$$(x - a)^2 + (y - b)^2 = r^2$$

表示为

$$x^2 + y^2 - 2ax - 2by = r^2 - a^2 - b^2.$$

由此变换为曲线

$$\frac{1}{\xi^2 + \eta^2} - \frac{2a\xi}{\xi^2 + \eta^2} - \frac{2by}{\xi^2 + \eta^2} = r^2 - a^2 - b^2$$

或

$$(\xi^2 + \eta^2)(r^2 - a^2 - b^2) + 2a\xi + 2b\eta = 1.$$

除非原来的圆通过原点, 这就表示 ξ, η 平面上的一个圆; 此外, 如果 $r^2 - a^2 - b^2 = 0$ 则像是一条直线.

(c) $-1/(x^2 + y^2)^2$.

5. 根据练习 4(b) 的解, 反变换映 $p_1p_2p_3$ 为对应角相等的通常的三角形.

6. 设 m_1 和 m_2 是通过点 (a,b) 的曲线斜率而 μ_1,μ_2 是它们的像的对应斜率. 根据

$$\mu=\frac{dv/dx}{du/dx}=\frac{\psi_x+m\psi_y}{\phi_x+m\phi_y}=\frac{\psi_x+m\psi_y}{\psi_y-m\psi_x},$$

立即得出

$$\frac{\mu_2-\mu_1}{1+\mu_2\mu_1}=\frac{m_2-m_1}{1+m_1m_2}.$$

7. 法线是

$$\frac{\xi-x}{u_x}=\frac{\eta-y}{u_y}=u-z.$$

当且仅当 $xu_y-yu_x=0$ 时它通过 z 轴. 曲面是旋转曲面的充分必要条件是 $z=f(w)$, 其中 $w=x^2+y^2$. 因此, 曲线 $z=$ 常数和 $w=$ 常数是同样的, 并且映像 $(x,y)\to(w,z)$ 的雅可比式必须为 0, 也就是,

$$\frac{d(w,z)}{d(x,y)}=2\begin{vmatrix} x & y\\ u_x & u_y\end{vmatrix}=0.$$

8. (a) 如果不是 $t<b$(椭圆) 就是 $b<t<a$(双曲线), 那么焦点是 $(0,\pm c)$, 其中 $c=\sqrt{a-b}$.

(b) 如果我们用 $F(x,y,t)$ 表示定义 t_1 和 t_2 的方程的左边, 则两条曲线 $t_1=$ 常数和 $t_2=$ 常数分别由隐函数方程

$$F(x,y,t_1)=1 \quad 和 \quad F(x,y,t_2)=1$$

给定. 因此, 它们正交的条件是

$$\begin{aligned}0&=F_x(x,y,t_1)F_x(x,y,t_2)+F_y(x,y,t_1)F_y(x,y,t_2)\\&=\frac{4x^2}{(a-t_1)(a-t_2)}+\frac{4y^2}{(b-t_1)(b-t_2)};\end{aligned}$$

但是这个关系是 $F(x,y,t_1)-F(x,y,t_2)=0$ 的直接结果.

(c) 用来定义 t_1 和 t_2 的二次方程的系数分别等于 t_1,t_2 和 (t_1+t_2). 于是我们得到两个关于 x^2 和 y^2 的线性方程, 从而

$$x=\pm\sqrt{\frac{(a-t_1)(a-t_2)}{a-b}},\quad y=\pm\sqrt{\frac{(b-t_1)(b-t_2)}{b-a}}.$$

(d) $\dfrac{d(t_1,t_2)}{d(x,y)}=\dfrac{4xy(a-b)}{\sqrt{\{(a+b)^2-2(a-b)(x^2-y^2)+(x^2+y^2)^2\}}}$.

(e) $\dfrac{f_1'g_1'}{(a-t_1)(b-t_1)}=\dfrac{f_2'g_2'}{(a-t_2)(b-t_2)}$.

9. (a) 设定义 t 的方程的左边是 $F(t)$. F 是 t 的连续函数 $(-\infty<t<c)$, 并且 $F(-\infty)=0$, $F(c-0)=+\infty$; 所以在这个区间内至少有一个点使得 $F=1$. 类似的结果也适用于别的区间.

(b) 参看练习第 8 题 (b).

(c) 参看练习第 8 题 (c),

$$x=\pm\sqrt{\frac{(a-t_1)(a-t_2)(a-t_3)}{(a-b)(a-c)}},$$

对于 y 和 z 有类似的公式.

10. (a) 应用练习第 6 题的结果.

(b) 设 $x=r\cos\theta$, $y=r\sin\theta$, 则直线 $\theta=$ 常数被变换到圆锥曲线 $t_1=\dfrac{1}{2}-\cos^2\theta$, 圆周 $r=$ 常数被变换到圆锥曲线

$$t_2=-\frac{1}{4}[r^2+(1/r^2)].$$

11. (b) 用 (24d) 如下:

$$x_{\xi\eta}=\frac{\partial}{\partial\xi}\left(\frac{-\xi_y}{D}\right)=x_{\eta\xi}=\frac{\partial}{\partial\eta}\left(\frac{\eta_y}{D}\right),$$

或者用 (a) 分题的结果.

练习 3.3 e (p.284)

1. (a) 1, (b) $4x^3$, (c) $\dfrac{\exp[2x/(x^2+y^2)]}{(x^2+y^2)^2}$.

2. (a), (c), 在 (b) 分题中, $u_0=v_0=1$ 不在复合变换的变化范围内.

3. 应用 (31b).

4. 反变换

$$x=p(\xi,\eta),\qquad y=q(\xi,\eta)$$

存在. 第一个结果这样获得: 通过与给定映像形成复合

$$z = f(p(\xi), q(\eta)) = \alpha(\xi, \eta),$$

$$\eta = \eta = \beta(\xi, \eta),$$

从而

$$\frac{d(z,\eta)}{d(\xi,\eta)} = \frac{d(z,\eta)}{d(x,y)} \cdot \frac{d(x,y)}{d(\xi,\eta)} = \frac{d(z,\eta)/d(x,y)}{d(\xi,\eta)/d(x,y)};$$

但是

$$\frac{d(z,\eta)}{d(\xi,\eta)} = \begin{vmatrix} \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} \\ 0 & 1 \end{vmatrix} = \frac{\partial z}{\partial \xi}.$$

练习 3.3 f (p291)

1. (a), (b). 在分题 (c) 中给定值不满足方程.

练习 3.3 g (p.298)

1. 在 $w = v - 1$ 情况下,

$$x_2 = 1 + \frac{1}{2}(u + w) + \frac{1}{8}(u^2 - 2uw - w^2),$$

$$y_2 = 1 - \frac{1}{2}(u - w) + \frac{1}{8}(u^2 + 2uw - w^2).$$

2. 相同.

练习 3.3 h (p.301)

1. $\xi = x^2 + x|x|, \eta = y$.
2. 若函数是相关的, 则 $\partial(\xi,\eta)/\partial(x,y) = a\beta - b\alpha = 0$.

练习 3.3 i (p.303)

1. (a) $-e^{3x}\cos y$,

   (b) 0,

(c) $-\left[\frac{y^z\log y\sinh x}{\cosh^2 y}-\frac{\cosh z}{y}-(\cosh z)y^{z-1}\sinh x\right]$,

(d) $-x^2\sin z$,

(e) x.

2. 存在这样一个区域，在其上 ξ,η,ζ 的某个函数恒等于零，条件是 $\partial(\xi,\eta,\zeta)/\partial(x,y,z)=0$.

3. 练习第 1 题 (b) 是相关的三元组：

$$(\eta^2+\rho^2)[(\eta+\rho-\xi)^2+\xi^2]=\alpha(\eta+\rho)^2.$$

4. $\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}=\begin{vmatrix}1 & 1 & 1\\ 2x & 2y & 2z\\ y+z & x+z & y+x\end{vmatrix}\equiv 0;\ \xi^2-\eta-2\zeta=0.$

5. (a) 因为两个曲面之间的夹角是它们法线之间的夹角，故我们只需证明任何两个方向之间的夹角是不变的. 设 s 是在 (x,y,z) 空间中的任何一条曲线的弧长，又 $t=(\dot{x},\dot{y},\dot{z})=\dot{\mathbf{X}}$ 是单位切向量，这里打点表示对 s 的微商. 显然 t 的方向映像到 τ 的方向，

$$\tau=\frac{(\dot{\xi},\dot{\eta},\dot{\zeta})}{(\dot{\xi}^2+\dot{\eta}^2+\dot{\zeta}^2)^{1/2}}=\dot{\mathbf{Y}}/|\dot{\mathbf{Y}}|,$$

像的方向 τ 通过 t 与 $\mathbf{X}$ 由下式给定

$$\tau=\mathbf{t}-\frac{2(\mathbf{t}\cdot\mathbf{X})\mathbf{X}}{|\mathbf{X}|^2}.$$

由此容易得出，在 $\mathbf{X}$ 相交的两条曲线之间的夹角的余弦由

$$\boldsymbol{\tau}_1\cdot\boldsymbol{\tau}_2=\mathbf{t}_1\cdot\mathbf{t}_2 \text{ 给定.}$$

(b) 仿效第 279 页练习第 4 题 (b) 的解答的做法.

(c) $-1/(x^2+y^2+z^2)^3$.

练习 3.4 a (p.314)

1. (a) $ds^2=\sin^2 v du^2+dv^2$,

(b) $ds^2 = \cosh^2 v du^2 + (1 + 2\sinh^2 v)dv^2$

(c) $ds^2 = (1 + f'^2)dz^2 + f^2 d\theta^2$,

$$
(d)\ ds^2 = \frac{(t_1 - t_2)(t_1 - t_3)}{4(a - t_1)(b - t_1)(c - t_1)} dt_1^2
$$

$$
+ \frac{(t_2 - t_1)(t_2 - t_3)}{4(a - t_2)(b - t_2)(c - t_2)} dt_2^2.
$$

2. $E = G = \cosh^2(t/a)$, $F = 0$.

3. $\mathbf{X}_u = (\cos V, \sin V, \alpha)$; $\mathbf{X}_v = (u\sin V, u\cos V, 0)$; 因此 $\mathbf{X}_u \cdot \mathbf{X}_v = 0$.

4. $ds^2 = (1 + z_x^2)dx^2 + 2z_x z_y dxdy + (1 + z_y)^2 dy^2$.

$$
5.\ EG - F^2 = \begin{vmatrix} y_u & z_u \\ y_v & z_v \end{vmatrix}^2 + \begin{vmatrix} z_u & x_u \\ z_v & x_v \end{vmatrix}^2 + \begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix}^2,
$$

用雅可比变换公式.

6. 引进坐标 x, y, z 使得 P 成为原点, P 点的切平面成为 xy 平面, t 成为 x 轴. 这样, S 的方程便取这样的形式 $z = f(x, y)$, 其中 $f(0,0) = f_x(0,0) = 0$. 经过 t 的一个平面 Σ 便由方程 $z = \alpha y$ 给定. 现在我们引进 $r = \sqrt{y^2 + z^2}$ 与 x 做为 Σ 上的坐标, 则 Σ 与 S 的交由隐式方程给出如下:

$$
\frac{r\alpha}{\sqrt{1 + \alpha^2}} = f\left\{x, \frac{r}{\sqrt{1 + \alpha^2}}\right\}.
$$

因此, 交线在点 $x = 0, r = 0$ 的曲率由下式给定 (参看第 241 页):

$$
k = f_{xx} \frac{\sqrt{1 + \alpha^2}}{\alpha}.
$$

于是, 这段的曲率中心的坐标是

$$
x = 0,\ y = \frac{1}{k\sqrt{1 + \alpha^2}} = \frac{\alpha}{f_{xx}(1 + \alpha^2)},
$$

$$
z = \frac{\alpha}{k\sqrt{1 + \alpha^2}} = \frac{\alpha^2}{f_{xx}(H\alpha^2)};
$$

也就是，它落在圆周 $f_{xx}(y^2+z^2)-z=0$ 上.

7. 取在 P 点的切平面作为 (x,y) 平面，则 S 的方程可以取为 $z=f(x,y)$，而法平面则由方程 $x=\alpha y$ 给定. 在这个平面上取 $r=\sqrt{x^2+y^2}$ 与 z 作为坐标，便有

$$z=f\left\{\frac{\alpha r}{\sqrt{1+\alpha^2}},\frac{r}{\sqrt{1+\alpha^2}}\right\},$$

并且它在 $r=0$ 处的曲率是

$$k=f_{xx}(0,0)\frac{\alpha^2}{1+\alpha^2}+\alpha f_{xy}(0,0)\frac{\alpha}{1+\alpha^2}+f_{yy}(0,0)\frac{1}{1+\alpha^2};$$

长度为 $\frac{1}{\sqrt{k}}$ 的向量的末端沿着直线 t 移动，于是有坐标

$$x=\frac{\alpha}{\sqrt{1+\alpha^2}}\frac{1}{\sqrt{k}},\ y=\frac{1}{\sqrt{1+\alpha^2}}\frac{1}{\sqrt{k}},\ z=0;$$

这就是说，它落在如下二次曲线上：

$$x^2f_{xx}+2xyf_{xy}+y^2f_{yy}=1.$$

8. (a) 对曲线的参量 t 将两个方程微商，我们得到

$$xx'+yy'+zz'=0,\quad axx'+byy'+czz'=0.$$

根据这些关系我们能找到比例 $x':y':z'$，这也就是切线的方向. 若 (ξ,η,ζ) 是流动坐标，则切线方程是

$$(\xi-x):(\eta-y):(\zeta-z)=\frac{c-b}{x}:\frac{a-c}{y}:\frac{b-a}{z}.$$

(b) 借助曲线方程的第二次微商并应用 (a) 的结果，我们得到

$$\begin{aligned}xx''+yy''+zz''&=-(x'^2+y'^2+z'^2)\\&=\lambda\left\{\frac{(c-b)^2}{x^2}+\frac{(a-c)^2}{y^2}+\frac{(b-a)^2}{z^2}\right\},\end{aligned}$$

$$axx'' + byy'' + czz''$$

$$= \lambda\left\{\frac{a(c-b)^2}{x^2} + \frac{b(a-c)^2}{y^2} + \frac{c(b-a)^2}{z^2}\right\},$$

这里 λ 是比例因子. 消去 λ, 我们有

$$(xx'' + yy'' + zz'')\left\{\frac{a(c-b)^2}{x^2} + \frac{b(a-c)^2}{y^2} + \frac{c(b-a)^2}{z^2}\right\}$$

$$= (axx'' + byy'' + czz'')\left\{\frac{(c-b)^2}{x^2} + \frac{(a-c)^2}{y^2} + \frac{(b-a)^2}{z^2}\right\}.$$

如果我们用 x', y', z' 代替 x'', y'', z'', 这个关于 x'', y'', z'' 的线性方程保持成立, 因此, 如果我们用某些线性组合

$$\lambda x' + \mu x'', \quad \lambda y' + \mu y'', \ \lambda z' + \mu z''$$

分别代替 x'', y'', z'', 它们仍旧满足. 现在如果 (ξ, η, ζ) 是在平面上, 则 $\xi - x, \eta - y, \zeta - z$ 正好是这样一个线性组合. (参看练习第 6 题, 第 232 页)

因此, 密切平面的方程是

$$\frac{ax^3}{c-b}(\xi - x) + \frac{by^3}{a-c}(\eta - y) + \frac{cz^3}{c-a}(\zeta - z) = 0.$$

9. 对两条曲线取 θ 作为参量. 则可在 (48) 中, 取 $u = \theta, v = \varphi$; 并且 $du/dt = dv/d\tau = 1$, $dv/dt = -1$, $dv/d\tau = 1$, $E = a^2, G = a^2\sin^2\theta$. 曲线的切线便由坐标向量 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 给定如下:

$$\dot{\mathbf{X}} = \mathbf{X}_\theta \pm \mathbf{X}_\phi = a(\cos\theta\cos\phi \pm \sin\theta\sin\phi)\mathbf{i}$$

$$+ a(\cos\theta\sin\varphi \mp \sin\theta\cos\phi)\mathbf{j} - a\sin\theta\mathbf{k},$$

并且 $|\dot{\mathbf{X}}|^2 = a^2(1 + \sin^2\theta)$ 在这两种情况下都一样; 还有

$$\ddot{\mathbf{X}} = 2a(\pm\cos\theta\sin\phi - \sin\theta\cos\phi)\mathbf{i}$$

$$+ 2a(\mp\cos\theta\cos\phi - \sin\theta\sin\phi)\mathbf{j} - a\cos\theta\mathbf{k}.$$

应用第 2.5 节练习第 8 题的公式便得结论.

练习 3.4 b (p.317)

1. 这映像除了在 $u = v = 0$ 外到处是共形的, 因为柯西 - 黎曼方程被满足. 在原点, 所有一阶微商都等于 0. 在极坐标

$$u = r\cos\theta, \quad v = r\sin\theta$$

中, 这映像成为 $x = r^2\cos 2\theta$, $y = r^2\sin 2\theta$; 于是, 在原点所有幅角都是成倍的.

2. 无论何时它是有定义的; 这也就是除了直线 $u = 0$ 以外到处都可以.

3. 以 $p = x\xi - y\eta$, $q = x\eta + y\xi$ 验证柯西 - 黎曼方程

$$\begin{aligned}\frac{\partial p}{\partial u} &= x\frac{\partial \xi}{\partial u} + \xi\frac{\partial x}{\partial u} - y\frac{\partial \eta}{\partial u} - \eta\frac{\partial y}{\partial u} \\ &= x\frac{\partial \eta}{\partial v} + \xi\frac{\partial y}{\partial v} + y\frac{\partial \xi}{\partial v} + \eta\frac{\partial x}{\partial v} \\ &= \frac{\partial q}{\partial v}.\end{aligned}$$

4. (a) 由 (40f) 推知

$$\mathbf{X}_u \cdot \mathbf{X}_u = \mathbf{X}_v \cdot \mathbf{X}_v = 4r^4/(u^2 + v^2 + r^2)^2,$$

$$\mathbf{X}_u \cdot \mathbf{X}_v = 0,$$

在 (48) 中令 $E = G$ 和 $F = 0$ 便得到所要求的结果.

(b) 球面上的圆周是球面与平面 (比如说 P) 的交线. 如果平面 P 经过北极, 测地投影就把圆周映射到 P 与 xy 平面的交线. 更一般地, 如果 P 有方程 $ax + by + cz = d$, 则根据 (40f) 有

$$(c - d)(u^2 + w^2) + 2ar^2u + 2br^2v = r^2(cr + d),$$

这在 $c=d$ 时是一个直线方程，而在 $c\neq d$ 时是一个圆方程.

(c) 根据 (40f),

$$u=x\left(1-\frac{z}{r}\right);\quad v=y\left(1-\frac{z}{r}\right).$$

对赤道平面的反射产生变换 $(u,v)\to(\xi,\eta)$, 这里

$$\xi=\frac{x}{1+z/r};\quad \eta=\frac{y}{1+z/r}.$$

从 (40f) 中换掉 x,y 与 z, 我们得到

$$\xi=\frac{r^2u}{u^2+v^2};\quad \eta=\frac{r^2v}{u^2+v^2}.$$

这是对半径为 r 的一个圆周的反演的方程组.

(d) 根据分题 (a) 的结果,

$$ds^2=\frac{4r^4}{(u^2+v^2+r^2)^2}(du^2+dv^2).$$

5. 按照 (48) 给定的角度必须满足

$$\cos\omega=\frac{du/dt\cdot du/d\tau+dv/dt\ dv/d\tau}{\sqrt{[(du/dt)^2+(dv/dt)^2][(du/d\tau)^2+(dv/d\tau)^2]}}.$$

取正交向量对 $(du/dt,dv/dt)=(0,1)$ 与 $(du/d\tau,dv/d\tau)=(1,0)$ 给出 $F=0$. 类似地，正交对 (1,1), (1,−1) 给出 $E=G$. 若 E 和 G 不是 0, 则满足条件

$$E=G,\qquad F=0.$$

6. 根据练习第 5 题的解答，我们要求

$$E=\sin^2\phi=\phi'^2=G.$$

解方程 $\phi'=\sin\phi$, 我们得到

$$v=\log\tan\frac{\phi}{2}\quad 或 \quad \phi=2\arctan e^v.$$

练习 3.5 a (p.320)

1. (a) 一族相似的椭圆，这些椭圆中心在原点，对称轴与坐标轴共线.

(b) 中心在 y 轴上并与 x 轴相切的圆族.

(c) 不是一族； c 的各个值产生同一条曲线，即单位圆

$$x^2 + y^2 = 1.$$

2. 半径为 1, 中心在如下直线上的球面:

$$x = y - 1 = \frac{1}{2}(z + \sqrt{2}).$$

练习 3.5 b (p.323)

1. 没有. 例如，可考虑一条直线或一个圆周的全部法线.

2. 一个包络，它满足参量方程

$$x = -\psi'(c), \quad y = -c\psi'(c) + \psi(c).$$

如果 ψ' 有一个逆 ϕ, 我们可以令 $\phi(-x) = (\psi')^{-1}(-x)$ 并利用 $c = \phi(-x)$ 来得到非参量方程

$$y = x\phi(-x) + \psi(\phi(-x)).$$

据此,

$$\begin{aligned} y' &= \phi(-x) - x'\phi'(-x) - \psi'(\phi(-x))\phi'(-x) \\ &= \phi(-x). \end{aligned}$$

在 y 的表达式中写入 $c = \phi(-x) = y'$, 就得到所要求的结果.

练习 3.5 c (p.330)

1. (a) 消去 t 得到

$$y = x\tan\alpha - \frac{g}{2v^2}x^2(1+\tan^2\alpha).$$

设 $c=\tan\alpha$ 是这个族的参量如下：

$$y = cx - \frac{1+c^2}{2v^2}gx^2, \tag{α}$$

其包络具有方程

$$y = \frac{v^2}{2g} - \frac{gx^2}{2v^2}.$$

(b) 对一个固定的 x 有

$$dy/dc = x - cgx^2/v^2, \quad d^2y/dc^2 = -gx^2/v^2 < 0.$$

因为在包络上，$dy/dc=0$, 我们由此得出结论：对任何一个给定的 x, 包络上的点是所能到达的目标的最高点.

(c) 对于 y 低于最大值的点 (x,y), 二次方程 (α) 对于 c 有两个解.

2. (a) 抛物线 $y^2=4x$;

(b) 直线 $x=\pm 2y$;

(c) 双曲线 $xy=\pm\frac{1}{2}$;

(d) 直线 $y=\pm ax$.

3. 设曲线的方程通过参量形式 $x=\varphi(t)$, $y=\psi(t)$ 给定. 圆族的包络满足

$$[x-\phi(t)]^2+[y-\psi(t)]^2=p^2,$$

$$[x-\phi(t)]\phi'(t)+[y-\psi(t)]\psi'(t)=0.$$

这是使动点 (x,y) 落在法线方向上并与点 $(\phi(t),\psi(t))$ 距离为 p 的精确条件.

4. 在曲线上我们可以引进 t 作为参量，使得这曲线通过

$$x=x(t),\ y=y(t),\ z=z(t)$$

给定，而且使得在具有参量 t 的点处的切线落在对应于 t 的两个平面上；这就给出这样两个关系：

$$ax'+by'+cz'=0,\quad dx'+ey'+fz'=0.$$

通过对 t 微商这些直线方程，就得到

$$a'x+b'y+c'z=0,\quad d'x+e'y+f'z=0.$$

又由关系

$$ax+by+cz=dx+ey+fz,$$

我们就有三个关于 x,y,z 的齐次方程而且系数行列式必须是 0.

5. (a) 以 t 为参量， C' 的参量方程组由如下方程组定义

$$\xi x+\eta y=1,\qquad \xi x'+\eta y'=0.$$

在第一个方程中，对 t 取通常微商，再用到第二个方程，我们就得到

$$\xi' x+\eta' y=0.$$

这与第一个方程结合起来，就确定了 C' 的极性逆，它显然就是曲线 C.

(b) $\xi^2(1-a^2)+\eta^2(1-b^2)-2ab\xi\eta+2a\xi+2b\eta=1.$

(c) $a^2\xi^2+b^2\eta^2=1.$

6. 母切线的方程是

$$x\sin\theta+y\cos\theta=a(\theta\sin\theta+\cos\theta-1).$$

7. 若 $(x^2/a^2)\pm(y^2/b^2)=1$ 是该圆锥曲线的方程，则

$$(x^2+y^2)^2=4(a^2x^2+b^2y^2)$$

是包络的方程. 注意，如果这圆锥曲线是矩形的双曲线，则这个包络就是普通双纽线 $(x^2+y^2)^2=4a^2(x^2-y^2)$.

8. (a) 若 Γ 是被参量形式 $\mathbf{X} = \Phi(t)$ 给定，则垂足曲线上的 $\mathbf{Y}$ 就由下列条件确定：

$$(\mathbf{Y} - \mathbf{X}) \cdot \mathbf{Y} = 0, \qquad \mathbf{Y} \cdot \mathbf{X}' = 0.$$

圆周上的点 Z 必须满足

$$\left(\mathbf{Z} - \frac{1}{2}\mathbf{X}\right)^2 = \frac{1}{4}\mathbf{X}^2 \text{ 或 } \mathbf{Z}^2 - \mathbf{Z} \cdot \mathbf{X} = 0.$$

于是为了在包络上 $\mathbf{Z}$ 必须满足 $\mathbf{Z} \cdot \mathbf{X}' = 0$. 这就是 $\mathbf{Z}$ 在垂足曲线上的条件.

(b) 从垂足曲线的原始定义，得一心脏线 $r = a(1 + \cos\theta)$, 其中 a 是圆周的半径, θ 是相对于从 0 到圆心方向的方位角.

9. 如果该椭圆有方程 $(x^2/a^2) + (y^2/b^2) = 1$, 则包络也是一个椭圆，其方程是

$$\frac{u^2}{b^2(a^2 + b^2)} + \frac{v^2}{b^2} = 1.$$

练习 3.5 d (p.334)

1. 这是一族椭球 $(x^2/a^2) + (y^2/b^2) + (z^2/c^2) = 1$, 满足

$$abc = k,$$

k 为一固定数. 其包络是

$$xyz = k^{2/3}\sqrt{27}.$$

2. 这是一族到原点具有单位距离的平面，其包络是单位球

$$x^2 + y^2 + z^2 = 1.$$

3. (a) $\sqrt{x} + \sqrt{y} + \sqrt{z} = 1$,

(b) $x^{2/3} + y^{2/3} + z^{2/3} = 1$.

4. 对包络我们有两个方程

$$x\cos t + y\sin t + z = t,$$

$$-x\sin t + y\cos t = 1.$$

这两个方程给出以 t 为参量的一族直线；如果以这样的直线作为切线的曲线存在，它也必须满足经过再一次微商而得到的方程.

(a) $r\sin[z+\sqrt{r^2-1}-\theta]+1=0$,

(b) 曲线给定由 $z=\theta-\pi/2$, $r=1$.

5. 设 $P(x,y,z)$ 是在管状曲面 Σ 上的一点，又设 S 是与 Σ 相交于 P 的一族球面. 于是 S 与 Σ 在 P 点有相同的切平面，这也就是说在这个点 x, y, z, z_x, z_y 有相同的值. 因此，只需证明，对于任何一个半径为 1 而中心在 xy 平面上的球面 (也就是对于 $u(x,y)=\sqrt{1-(x-a)^2-(y-b)^2}$) 这样的关系成立.

6. 利用反演. 因为 S_1, S_2, S_3 经过原点，它们被变换成平面；所以我们只需求得这样一些球面的包络，这些球面与三个平面相切 (也就是某个圆锥)，我们再把这包络反演为

$$(x^2+y^2+z^2)^2-2(x^2+y^2+z^2)(x+y+z)$$

$$-3(x^2+y^2+z^2-2xy-2xz-2yz)=0.$$

7. (a) 若 P 描成 Γ 的垂足曲线 Γ'，在垂直于 Γ 所在平面的平面上，以 OP 为直径作一个圆；则包络是通过这些变动的圆周所产生的曲面.

(b) 见分题 (a) 与第 3.5 节 c 的练习 8(b) 的解答.

8. 这是一族平面 $(x/a)+(y/b)+(z/c)=1$，满足 $abc=K$；包络则是通过这个方程连同下列两个方程来确定：

$$-\frac{x}{a^2}+\frac{zK}{c^2a^2b}=0; \qquad -\frac{y}{b^2}+\frac{zK}{c^2ab^2}=0.$$

这结合第一个方程便给出

$$x/a=y/b=z/c=\frac{1}{3},$$

从而

$$xyz = \frac{K}{27}.$$

9. 这样一个平面必须包含第一条抛物线在点 $(a^2, 2a, 0)$ 处的切向量 $T_1 = (a, 1, 0)$, 还包含第二条抛物线在点 $(b^2, 0, 2b)$ 处的切向量 $T_2 = (b, 0, 1)$. 切线相交的条件给出 $b = +a$, 以及交点 $(-a^2, 0, 0)$. 用 $T_1 \times T_2 = (1, -a, -b)$ 作为平面的法线, 我们就得到它的方程 $x - a(y+z) + a^2 = 0$, 以 a 作为参量, 而以抛物柱面 $4x = (y+z)^2$ 作为包络.

练习 3.6 a (p.339)

1. (a) $-\sin v$,

   (b) $(a^3 + b^3 + c^3)(u - v) + 3abcv$,

   (c) $4uv$.

练习 3.6 b (p.341)

1. (a) $-2xydxdy$;

   (b) $(x^4 - 4x^2y^2 + y^4)dxdy$;

   (c) $(a^2 + b^2)dxdydz$.

2. 因为 $\omega = Adx + Bdy + Cdz$, 我们有

$$\omega^2 = A^2dxdx + B^2dydy + C^2dzdz + AB(dxdy + dydx)$$

$$+ BC(dydz + dzdy) + CA(dzdx + dxdz),$$

而在 ω^2 中的每一项都等于零. 换句话说, 因为我们知道对于任何两个这样的型使得 $\omega_1\omega_2 = -\omega_2\omega_1$ 的都有 $\omega^2 = -\omega^2$; 所以 $\omega^2 = 0$.

3. 应用练习第 2 题的结果.

4. 改写左边成如下形式

$$[(\omega_1 + \omega_3) + (\omega_2 + \omega_4)][(\omega_1 + \omega_3) - (\omega_2 + \omega_4)],$$

再运用练习第 3 题的结果.

5. $L_1(L_2L_3)=(A_1dx+B_1dy+C_1dz)$

$$\times\left\{\begin{vmatrix}B_2 & B_3\\ C_2 & C_3\end{vmatrix}dydz+\begin{vmatrix}C_2 & C_3\\ A_2 & A_3\end{vmatrix}dzdx+\begin{vmatrix}A_2 & A_3\\ B_2 & B_3\end{vmatrix}dxdy\right\}$$

$$=\left\{A_1\begin{vmatrix}B_2 & B_3\\ C_2 & C_3\end{vmatrix}+B_1\begin{vmatrix}C_2 & C_3\\ A_2 & A_3\end{vmatrix}+C_1\begin{vmatrix}A_2 & A_3\\ B_2 & B_3\end{vmatrix}\right\}dxdydz,$$

其中 $dxdydz$ 的系数是对于行列式

$$\begin{vmatrix}A_1 & B_1 & C_1\\ A_2 & B_2 & C_2\\ A_3 & B_3 & C_3\end{vmatrix}$$

按第一行的子式展开式.

练习 3.6 c (p.346)

1. (a) $-\dfrac{y}{x^2+y^2}dx+\dfrac{x}{x^2+y^2}dy,$

(b) $2dxdy,$

(c) $0,$

(d) $x(\cos y-1)\sin z,$

(e) $0.$

2. 因 $\omega_i=A_idx+B_idy+C_idz(i=1,2)$, 我们有

$$d(\omega_1\omega_2)=\left\{\left(\frac{\partial B_1}{\partial x}C_2+B_1\frac{\partial C_2}{\partial x}-\frac{\partial C_1}{\partial x}B_2-C_1\frac{\partial B_2}{\partial x}\right)\right.$$

$$+\left(\frac{\partial C_1}{\partial y}A_2+C_1\frac{\partial A_2}{\partial y}-\frac{\partial A_1}{\partial y}C_2-A_1\frac{\partial C_2}{\partial y}\right)$$

$$+\left(\frac{\partial A_1}{\partial z}B_1+A_1\frac{\partial B_2}{\partial z}-\frac{\partial B_2}{\partial z}A_2\right.$$

$$\left.\left.-B_1\frac{\partial A_2}{\partial z}\right)\right\}dxdydz$$

$$= \left\{ \left(\frac{\partial C_1}{\partial y} - \frac{\partial B_1}{\partial z} \right) A_2 + \left(\frac{\partial A_1}{\partial z} - \frac{\partial C_1}{\partial x} \right) B_2 \right.$$

$$\left. + \left(\frac{\partial B_1}{\partial x} - \frac{\partial A_1}{\partial y} \right) C_2 \right\} dxdydz$$

$$+ \left\{ A_1 \left(\frac{\partial B_2}{\partial z} - \frac{\partial C_2}{\partial y} \right) + B_1 \left(\frac{\partial C_2}{\partial x} - \frac{\partial A_1}{\partial z} \right) \right.$$

$$\left. + C_1 \left(\frac{\partial A_2}{\partial y} - \frac{\partial B_2}{\partial x} \right) \right\} dxdydz$$

$$= (d\omega_1)\omega_1 + \omega_1(d\omega_2).$$

3. 根据练习第 2 题，若 $d\omega_1 = d\omega_2 = 0$, 则 $d(\omega_1\omega_2) = 0$.

练习 3.6 d (p.356)

1. 考虑 $F(\mathbf{X}) = f(\rho, \phi, \theta) = g(x, y, z)$ 作为空间的点的函数，根据微分的形式不变性我们知道

$$dF = dg = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy + \frac{\partial g}{\partial z} dz = \nabla F \cdot d\mathbf{X}$$

$$= \frac{\partial f}{\partial \rho} d\rho + \frac{\partial f}{\partial \phi} d\phi + \frac{df}{d\theta} d\theta.$$

因此，

$$\nabla F \cdot d\mathbf{X} = \left(\frac{\partial f}{\partial \rho} \mathbf{u} + \frac{1}{\rho} \frac{\partial f}{\partial \phi} \mathbf{v} + \frac{1}{\rho \sin \varphi} \frac{\partial f}{\partial \theta} \mathbf{w} \right) \cdot d\mathbf{X},$$

从而

$$\nabla F = \frac{\partial f}{\partial \rho} \mathbf{u} + \frac{1}{\rho} \frac{\partial f}{\partial \phi} \mathbf{v} + \frac{1}{\rho \sin \phi} \frac{\partial f}{\partial \theta} \mathbf{w},$$

练习 3.7 b (p.362)

1. (a) 鞍点在

$$y = 0, x = \pi/3 + 2n\pi;$$

极小点在

$$y = 0, \quad x = -\pi/3 + 2n\pi.$$

(b) 极大在 $x=\pi/4+2n\pi$, $y=\pi/4+2n\pi$ 和 $x=3\pi/4+2n\pi$, $y=3\pi/4+2n\pi$; 极小在 $x=\pi/4+2n\pi$, $y=3\pi/4+2n\pi$ 和 $x=3\pi/4+2n\pi$, $y=\pi/4+2n\pi$.

(c) 鞍点在 $x=0, y=1$.

(d) 没有稳定点.

(e) 鞍点在 $x=0$, $y=0$.

2. 极大在 $x=0, y=\pm1$; 极小在 $x=y=0$.

3. 极小在 $x=1, y=4$; 鞍点在 $x=-1, y=2$.

4. $a/20, a/10, a/10$.

5. 在平面上非正常极小点

$$x=0,\quad y=1,\quad z=-\frac{1}{2}.$$

6. 极大化 $V=xy[100-2(x+y)]$. 当 $x=y=50/3$, $z=100/3$ 时体积达到极大, $V_{\max}=(25/27)\times10^4$ 英寸 $^3\approx5.4$ 英尺 ^{3}.

7. 令 $\mathbf{X}=(x,y,z)$ 并设 n 个点是 $(a_i,b_i,c_i)(i=1,2,\cdots,n)$. 为了极小化 $\Sigma[(x-a_i)^2+(y-b_i)^2+(z-c_i)^2]$, 令

$$2\Sigma(x-a_i)=2\Sigma(y-b_i)=2\Sigma(z-c_i)=0;$$

因此 $x=(1/n)\Sigma a_i$, $y=(1/n)\Sigma b_i$, $z=(1/n)\Sigma c_i$. 该和数在 n 个点的重心处达到极小.

练习 3.7 c (p.367)

1. 取 $F(x,y,z)=xyz+\lambda[2(x+y)+z-100]$.
由 $F_x=yz+2\lambda,\quad F_y=zx+2\lambda,\quad F_z=xy+\lambda$ 知极值出现时有

$$V=xyz=-2\lambda x=-2\lambda y=-\lambda z,$$

因而, $z=2x=2y$. 将此写入约束条件, 就得到

$$z=100/3, x=y=50/3.\qquad \text{这与前面一样.}$$

2. $x = y = \frac{1}{2}, z = \frac{1}{16}$.

3. $x = -y = 1/\sqrt{2}, z = 1$.

4. 取这 n 个点的重心为原点，并设它们的坐标是 (a_i, b_i). 设 $\mathbf{X} = (x, y)$ 并设这直线由 $Ax + By = C$ 给定. 把拉格朗日乘数法运用到

$$\Sigma[(x - a_i)^2 + (y - b_i)^2] + (C - Ax - By),$$

我们就得到

$$2nx - \lambda A = 2ny - \lambda B = 0;$$

从而

$$\lambda = \frac{2nC}{A^2 + B^2}.$$

于是

$$x = \frac{AC}{A^2 + B^2}, \qquad y = \frac{BC}{A^2 + B^2};$$

这就是说，$\mathbf{X}$ 是在直线上到重心最近的点.

5. 设 S 表示曲线 $f(x, y) = C$, S' 表示曲线 $\phi(x, y) = C'$. S 与 S' 在 (a, b) 有接触点. 一般说来，在某个邻域内 $f(x, y) - C$ 在 S 的一侧是正的，而在 S 的另一侧是负的；这对于 $\phi(x, y) - C'$ 与 S' 也是类似的. 例如，如果 $f(a, b)$ 是 f 的极大值则在 S' 上 $f(x, y) - C \leq 0$, 也就是 S' 整个地在 S 的一边上，于是 S 也就是在 S' 的一边上. 这也就是说，$\phi(x, y) - C'$ 在 S 上有不变的符号，而当它在 (a, b) 等于 0 时，它就在该点处有一个极大或一个极小.

练习 3.7 e (p.375)

1. 对于光滑的 f 与 ϕ, 极小值 C 象征着等量面 $f(x, y, z) = C$ 与曲面 $\phi(x, y, z) = 0$ 相切.

2. 在两个柱面 $\varphi(x, y) = 0, \psi(x, z) = 0$ 的交线上寻找 $f(x, y, z)$ 的一个极值点. 假设 f 是光滑的而且交线是光滑曲线，那么这点就出现在 f 的等量面与曲线相切处.

练习 3.7 f (p.379)

1. 对下式求极值

$$(x-a)^2+(y-b)^2+(z-c)^2+\lambda(D-Ax-By-Cz)$$

得到条件

$$2(x-a)-\lambda A=2(y-b)-\lambda B=2(z-C)-\lambda C=0;$$

从而

$$\lambda=\frac{2(D-aA-bB-cC)}{A^2+B^2+C^2}.$$

这给出

$$x=a+\frac{A(D-aA-bB-cC)}{A^2+B^2+C^2},\cdots,$$

因而最小距离 p 给定如下:

$$p=\frac{|D-aA-bB-cC|}{\sqrt{A^2+B^2+C^2}}.$$

2. $(4+\sqrt{5})/\sqrt{2},(4-\sqrt{5})/\sqrt{2}$.

3. 最大值同对于表达式 $ax^2+2bxy+cy^2$ 在辅助条件

$$ex^2+2fxy+gy^2=1$$

之下求极大值是相同的.

4. 参看练习第 3 题.

(a) $14/3+2\sqrt{67}/3$,

(b) 函数有一个非严格的极大值 (p.357) 等于 1.95, 当 $y/x=0.64$ 时.

5. 椭圆显然与圆相切; 也就是说, 这两个方程必须对 x 给出两重根. 因此, 接触条件是

$$a^2(b^2-1)=b^4,a=3/\sqrt{2},b=\sqrt{3/2}.$$

6. $(-1/\sqrt{14}, -2/\sqrt{14}, -3/\sqrt{14})$. 这是在联接给定点与中心的直线上.

7. $A = a^2/x$, $B = b^2/y$, $C = c^2/z$, 以及如下辅助条件一起 $(x^2/a^2) + (y^2/b^2) + (z^2/c^2) = 1$.

(a) $x = \dfrac{a^{4/3}}{\sqrt{a^{2/3} + b^{2/3} + c^{2/3}}}, \cdots$,

(b) $x = \dfrac{a^{3/2}}{\sqrt{a + b + c}}, \cdots$.

8. 顶点给定于 $x = \pm a/\sqrt{3}, y = \pm b/\sqrt{3}, z = c/\sqrt{3}$.

9. 顶点给定于 $x = a^2/\sqrt{a^2 + b^2}, y = b^2/\sqrt{a^2 + b^2}$.

10. $x = 1, y = 1$.

11. 最长轴通过 $\sqrt{x^2 + y^2 + z^2}$ 的最大值给定，其中辅助条件是 (x, y, z) 落在椭球上. 因此，我们有这样三个方程

$$\frac{x}{\sqrt{x^2 + y^2 + z^2}} = \frac{x}{l} = \lambda(ax + dy + ez), \cdots$$

分别用 x, y, z 乘这些等式而后相加，我们有

$$\lambda = \sqrt{x^2 + y^2 + z^2} = l.$$

另一方面，我们可以把这些方程看作是三个关于 x, y, z 的线性方程组，它们的行列式必须等于零.

12. (a) 相当于对下式求极大

$$a \log x + b \log y + c \log z + \lambda(1 - x^k - y^k - z^k).$$

这给出

$$\lambda x^k = \frac{a}{k}, \lambda y^k = \frac{b}{k}, \lambda z^k = \frac{c}{k};$$

从而

$$\lambda = \frac{1}{k}(a + b + c).$$

当

$$x^k = \frac{a}{a + b + c}, \quad y^k = \frac{b}{a + b + c}, \quad z^k = \frac{c}{a + b + c},$$

最大值被达到而且等于

$$k\sqrt{\frac{a^a b^b c^c}{(a+b+c)^{a+b+c}}}.$$

(b) 设 $x^k = u/(u+v+w)$, $y^k = v/(u+v+w)$, $z^k = w/(u+v+w)$, 并代入

$$(x^a y^b z^c)^k \leq \frac{a^a b^b c^c}{(a+b+c)^{a+b+c}}.$$

13. 参看第 338 页上对三角形的类似的证明. 极小点 0 确实存在. 首先证明, 0 不是顶点之一, 于是它只能是对角线的交点. 用这样一个事实, 向量和为 O 的 4 个单位向量的末点形成一个矩形. 然后证明, 对角线交点与任何一个顶点相比, 它到顶点的距离之和是较小的.

14. 假设 a, b 这一对与 c, d 这一对是相邻的. 设 ϕ 是 a 与 b 之间的夹角, ψ 是 c 与 d 之间的夹角. 问题是要在条件

$$f(\varphi, \psi) = (a^2 + b^2 - 2ab\cos\varphi) - (c^2 + d^2 - 2cd\cos\psi) = 0$$

之下使

$$A(\varphi, \psi) = \frac{1}{2}(ab\sin\varphi + cd\sin\psi)$$

最大. 让各个微商 $(\partial/\partial\varphi)(A+\lambda f)$ 与 $(\partial/\partial\psi)(A+\lambda f)$ 等于 0, 我们得到

$$\lambda = -\frac{1}{4\tan\varphi} = \frac{1}{4\tan\psi},$$

从而 $\varphi + \psi = \pi$, 因此,

$$A = \frac{1}{2}(ab + cd)\sin\varphi,$$

这里

$$\cos\varphi = \frac{1}{2}(a^2 + b^2 - c^2 - d^2)/(ab + cd).$$

消去 ϕ, 我们得到最大面积

$$A = \frac{1}{4}\sqrt{4(ab+cd)^2 - (a^2+b^2-c^2-d^2)}$$
$$= \frac{1}{4}\sqrt{8abcd - (a^2+b^2+c^2+d^2)^2}.$$

它显然与我们关于边的顺序的假定无关.

最大值与边的顺序无关这个结论在几何上是显然的, 因为任何一对相邻边可以交换而不影响凸多边形的面积.

练习 A.1 (p.387)

1. (a) 极小在原点.

(b) 为简单起见, 引进新变量 $u = x + y$, $v = x - y$. 我们求下式的极值:

$$f(u,v) = \cos u + \sin v + \frac{1}{4}(u+v)^2.$$

条件 $f_u = f_v = 0$ 给出

$$\cos v = -\sin u = -\frac{1}{2}(u+v). \qquad \text{(i)}$$

我们必须分下列两种可能情况:

1) $\sin v = -\cos u$. 在这种情况下 $f'_{uv} - f_{uu}f_{vv} = \cos^2 u$ 只能找到鞍点.

2) $\sin v = \cos u$. 在这种情况下, (i) 式给出

$$u + v = -\frac{\pi}{2},$$

我们可以有 $u = -\alpha$ 或 $u = \pi + \alpha$. 在前一种情况下,

$$f_{uv}^2 - f_{uu}f_{vv} = \cos u(1 - \cos u)$$

是正的, 因而得到鞍点; 在后一种情况, 它是负的, 因而由于

$$f_{uu} = f_{uv} = \cos\alpha + \frac{1}{2}$$

我们得到一个极小. (c) 无极值, 因为 $f_x > 0$ 处处成立.

2. $$\begin{aligned}f(x)+f(y)+f(z) &= 3f(a)+\{(x-a)+(y-a)\\&\quad+(z-a)\}f'(a)+\frac{1}{2}\rho^2\{f''(a)+\varepsilon\},\end{aligned}$$

其中 $\rho^2=(x-a)^2+(y-a)^2+(z-a)^2$. 另一方面, 辅助条件给出

$$\begin{aligned}&(x-a)+(y-a)+(z-a)\\&\quad=\rho^2\left(-\frac{\phi''(a)}{2\phi'(a)}+\varepsilon\right)-\frac{\phi'(a)}{\phi(a)}\{(x-a)(y-a)\\&\quad+(x-a)(z-a)+(y-a)(z-a)\}\\&\quad=\left(-\frac{\phi''(a)}{2\phi'(a)}+\frac{\phi'(a)}{2\phi(a)}+\varepsilon\right)\rho^2,\end{aligned}$$

其中 $\lim\limits_{x,y,z\to a}\varepsilon=0$.

3. 如果 $P_i=(x_i,y_i), r_i=PP_i$, 则我们有

$$d^2f=\sum_1^3 d^2r_i=\sum_{i=1}^3 r_i^{-3}[(y-y_i)dx-(x-x_i)dy]^2,$$

它是正定的.

4. 在点 P_1 处. 注意函数 $f=r_1+r_2+r_3$ 在全平面上是连续的, 但在点 P_1,P_2,P_3 不可微而有锥形点 (如像函数

$$z=\sqrt{(x-x_1)^2+(y-y_1)^2},$$

它在几何上表示一个圆锥). 在点 P_1 沿着围绕这点的各个方向上来研究函数 f 的微商.

5 (a) 如果我们表达

$$f=lx+my+nz, \psi=x^p+y^p+z^p-c^p, F=f-\lambda\phi,$$

则稳定值条件是

(1) $$l = \lambda p x^{p-1}, m = \lambda p y^{p-1}, n = \lambda p z^{p-1}.$$

用 x, y, z 分别乘这些方程然后相加，我们有

(2) $$lx + my + nz = \lambda p c^p.$$

根据 (1) 计算 x, y, z 并代入到 $\phi = 0$ 中去，我们得到

$$\lambda p = (l^q + m^q + n^q)^{1/q} c^{1-p}.$$

将 λp 的这个表达式代入 (2), 就给出稳定值.

(b) 参看练习第 6 题. 这里我们有

$$d^2 F = -\lambda p(p-1)(x^{p-2}dx^2 + y^{p-2}dy^2 + z^{p-2}dz^2);$$

当 $p > 0$ 时，这个二次型是正定的还是负定的，随 $p \gtrless 1$ 而定.

6. 这证明类似于在 $n = 2$ 情形 (p.357). 一个正定二次型 $\Sigma a_{ik} x_i x_k$, 在行列式不等于零的情况下，通过适当的变换

$$x_i = \sum_{k=1}^{n} c_{ik} y_k \qquad (i = 1, 2, \cdots n)$$

能被简化成形状

$$\Sigma a_{ik} x_i x_k = y_1^2 + y_2^2 + \cdots + y_n^2 > m(x_1^2 + \cdots + x_n^2),$$

其中 m 是一个适当的正常数. 对于应用来说，记住这样一个结论是重要的：一个二次型 $\Phi = \Sigma a_{ik} x_i x_k$ 为正定的充分必要条件是，它的如下所示的 $1, 2, \cdots, n$ 阶各个主子式

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & \vdots \\ a_{31} & a_{32} & a_{33} & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

都应当是正的；而 Φ 是负定的，如果 $-\Phi$ 是正定的.

7. 根据第一个法则, 我们应该从 (3) 计算 d^2f, 而由 (1) 代入 $dx_1,\cdots,dx_m,d^2x_1,\cdots,d^2x_m$. 注意 (1) 隐含着

$$d^2\phi_\mu=\sum\phi_{\mu x_i x_k}dx_i dx_k+\phi_{\mu x_1}d^2x_1+\cdots+\phi_{\mu x_m}d^2x_m=0$$

$$(\mu=1,2,\cdots,m).$$

如果这式乘以 $\lambda\mu$, 并对一切 μ 的值与 (3) 相加, 我们就有

$$d^2f=d^2F=\Sigma F_{x_i x_k}dx_i dx_k,$$

因为由于关系 (2), $d^2x_1,\cdots,d^2x_m$ 都消去了.

8. 对于 $F=f+\lambda\phi$(省去一个正因子), 我们得到

$$d^2F=\sum_{i,k=1,\cdots,n}dx_i dx_k$$

$$(d\phi=dx_1+\cdots+dx_n=0).$$

消去 dx_n, 我们需要证明的是, 二次型

$$-d^2F=(dx_1+\cdots+dx_{n-1})^2-\sum_{i,k=1,\cdots,n-1}dx_i dx_k$$

$$=\sum_{i=1,\cdots,n}dx_i^2+\sum_{i,k}^{1,n-1}dx_i dx_k$$

是正定的.

9. 由 $dx=-dy-dz$ 有

$$d^2F=-2s[(s-z)dy^2+(s-x)dydz+(s-y)dz^2].$$

当 $x=y=z$ 时 d^2F 的判别式是正的, 而 d^2F 是负定的.

练习 A.2(p.397)

1. (c) 用极坐标 $x=r\cos\theta$, $y=r\sin\theta$, 取

$$f(x,y)=r^{n+1}\sin(n+1)\theta,$$

对此有

$$\nabla f=(n+1)r^n(\sin n\theta,\cos n\theta).$$

3. 如果不存在不动点，则在 R 中到处都有 $u^2+v^2\neq 0$. 因为凸区域 R 是单连通的，可像第 396 页中那样推知，曲线 C 相对于向量场的指数 I_c 是 0. 另一方面，因为 R 被映射到自身，所以对于 C 上的每一个点，向量 (u,v) 指向到 R 内或切线方向. 这隐含着

$$I_c=1/2\pi\int_c d\theta=1,$$

如果 C 具有取决于坐标系的通常的定向值.

练习 A.3 (p.400)

1. (a) 结点在 (0,0), 具有切线 $x=\pm y$.

(b) 方程

$$f_x=2x-6x^2+4xy^2=0,$$

$$f_y=2y-6y^2+4x^2y=0$$

有公共解

$$(0,0),\left(\sqrt{\frac{1}{2}},0\right),\left(0,\sqrt{\frac{1}{2}}\right),\left(\frac{1}{2},\frac{1}{2}\right)\text{和}(1,1),$$

其中只有第一个与最后一个是曲线上的点. 在 (0,0), 奇点是一个孤立点; 在 (1,1), $f_{xx}=f_{yy}=0$ 且 f_{xy}=8; 奇点是一个结点，具有切线 $x=1$ 与 $y=1$.

(c) 在 (0,0) 点有双重切线 $y=x$. 曲线有两个分支; 准确到二阶 $y=x\pm x^2$.

(d) 在 (0,0) 点有双重切线. 曲线有一个尖点. 这就是第 3.2b 节，练习第 3 题的同一条曲线.

练习 A.4 (p.401)

1. 如果该二次型是非退化的而且是确定的，则它的奇点是一个孤立点；如果是非退化但是不定的，则切线在奇点处形成一个锥. 如果这二次型是退化与半定的，则切线可以落在一个平面上，在这个平面上两个分支互相相切，如像平面 $z=0$ 对于曲面

$$z^{2/3}+(x^2+y^2)^{2/3}=a^{2/3}$$

在 $(a,0,0)$(一个线型尖点)，如像

$$z^4=(x^2+y^2)^3$$

在 $(0,0,0)$(两个相切的分支). 或者可以有这样类型的尖点，在这样点上仅有一条切线存在，如像直线 $x=y=0$ 对于前一个曲面在 $(0,0,a)$ 一样. 如果这二次型是退化的而且是不定的，则切线落在两个平面上，如像平面 $x=\pm y$ 在 $(0,0,0)$ 对于

$$x^2-y^2+z^2=0$$

那样.

练习 A.5 (p.403)

1. 这个流动是稳定的，也就是说，在空间的每一个点处流速对时间是常数.

2. 若 $\mathbf{U}=(u,v,w)$ 是这质点在时刻 t 经过点

$$\mathbf{X}=(x,y,z)$$

时的速度，则它的加速度是

$$\begin{aligned}\frac{d^2\mathbf{X}}{dt^2}&=\frac{d\mathbf{U}}{dt}=\frac{d\mathbf{X}}{dt}\cdot\nabla\mathbf{U}+\frac{\partial\mathbf{U}}{\partial t}\\&=\mathbf{U}\cdot\nabla\mathbf{U}+\frac{\partial\mathbf{U}}{\partial t}.\end{aligned}$$

练习 A.6 (p.405)

1. (a) $x = -2 - 2\cos\alpha, y = -2\sin\alpha$ 或

   $(x+2)^2 + y^2 = 4;$

   $L = 4\pi, A = 4\pi.$

   (b) $x = -\sin^3\alpha, y = -\cos^3\alpha$ 或

   $x^{2/3} + y^{2/3} = 1,$

   $L = \frac{3}{2}\int_0^{2\pi} |\sin 2\alpha| d\alpha = 6\int_0^{2\pi} \sin 2\alpha d\alpha = 6.$

$A = -(3/8)\pi$, 这里负号来源于曲线的顺时针定向.

2. 是. 对于大的 c, 考虑直角三角形, 其顶点为 $(0,0)$, $(0,c)$, $(c^{-2},0)$.

3. 对于能够表示成它的切线包络的曲线, 它必须是分段光滑的.

数学名著译丛

拓扑空间论

代数特征值问题

数学概观

常微分方程

数学与猜想

代数几何

数学——它的内容，方法和意义

非线性与泛函分析

几何基础（第二版）

